Analysis of a DNA methyltransferase homologue in fission yeast

Caroline R. M. Wilkinson

Thesis presented for the Degree of

Doctor of Philosophy

at the University of Edinburgh

1994

Dedication

This thesis is dedicated to my grandma,
Dorothea Wilkinson.

"Well there's always retrospect, to light a clearer path.
Every five years or so I look back on my life and I have a good laugh."

Emily Saliers-"Watershed"

Acknowledgments

Well, things didn't exactly go according to plan and there have been some rather frustrating times over the last three years but I have managed to come out the other end with my sanity and sense of humour just about intact. This would not have been possible without the help and support of lots of people and I would like to thank:

-my supervisor Adrain Bird, for many stimulating and enlightening discussions and for valiantly doing battle with my appalling grammar (well I am from Yorkshire, what did you expect?)

-all the members of the Bird lab, past and present, for making it such a happy place to work and in particular,

-Richard Meehan, for constant technical advice, encouragement and confidence boosting throughout the last three years and for helpful comments on this thesis and life in general, all very much appreciated.

-the exquisite Ms. Jillian Charlton for lots of advice, technical and otherwise and much moral support.

- Sally Cross for lots of help in the lab and great games of squash.

-Susan Tweedie for RT-PCR advice and many helpful comments on the introduction and thesis writing in general.

-Paco Antequera for instilling in me the principles of classical science!

-all my office chums, especially Javier for his inspirational singing and his handy hints scheme !!!!!

-Joan Davidson and Aileen Grieg for their excellent technical assistance and for generally looking after me!!

-Christine Struthers for her wonderful organizational skills!

-to Peter Fantes for allowing me to use the *pombe* facilities in his lab and for useful advice.

-to the members of the Fantes lab, all of whom have given me an enormous amount of help and I would like to thank in particular,

-Chris McInerny for all his encouragement, especially on RNA matters.

-Stuart MacNeill, for lots of useful advice and helpful comments on various bits of this thesis,

-and the other one, what's his name, something Mackie, he's been quite helpful too.

-thanks to George for brightening up Friday afternoons

-to Frank Johnston and Graham Brown for their excellent photography service
-to David Hornby for useful discussions and for mutual winging sessions.
-to Elmar Maier, whose speedy analysis of the *pmt1* mapping data saved me from a lot more gratuitous sequencing!
-to Robin Allshire for the mini-chromosome and useful discussions.
-to Rachel Bartlett and Paul Nurse for providing me with such a character building exercise. nice one.
-and finally, thanks to my family and to Shaun for their love and support.

Contents

Preface

Chapter 1 Introduction

Chapter 2 Materials and Methods

Chapter 5 Purification of the cnd1 protein and analysis of its methyltransferase activity

Chapter 6 The *S. pombe* m^5C-MTase homologue is not the *cnd1*$^+$ gene

Chapter 7 Analysis of a *pmt1* null mutation

Chapter 8 Discussion

Abbreviations

Ab	Antibody
AdoMet	S-adenosyl-L-methionine
5-azaC	5-azacytidine
bp	base pairs
BSA	bovine serum albumin
ATP	adenosine triphosphate
cDNA	complementary DNA
cnd	completion of nuclear division
cpm	counts per minute
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
DEPC	diethyl pyrocarbonate
dGTP	deoxyguanosine triphosphate
DMSO	dimethyl sulphoxide
dNTPs	all four deoxy triphosphates
DTT	dithiothreitol
dTTP	deoxythymidine triphosphate
DNA	deoxyribonucleic acid
EDTA	ethylenediaminetetraacetic acid
EMM	Edinburgh minimal medium
ExoIII	Exonuclease III
h	hour
Hepes	N-(Hydroxyethyl)piperazine-N'-[2-ethanosulphonic acid]
IPTG	isopropylthio-β-D-galactoside
kb	kilobase
kD	kilodalton
LTR	long terminal repeat
m^5C	5-methylcytosine
m^5C-MTase	cytosine-5-specific DNA methyltransferase
MBD	methyl binding domain
MNE	mops sodium acetate EDTA
min	minutes
MOPS	3-(N-morpholino)propane-sulphonic acid
mRNA	messenger RNA
MW	molecular weight
OD_x	optical density at xnm
ORF	open reading frame
PCR	polymerase chain reaction
PEG	polyethylene glycol
PMSF	phenylmethylsulphonylfluoride
pmt	pombe methyltransferase
PNK	polynucleotide kinase
RNA	ribonucleic acid
rpm	revolutions per minute
RT	reverse transcriptase

scd	suppressor of cnd1
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
ssDNA	single-stranded DNA
TAE	Tris acetate EDTA
TBE	Tris borate EDTA
TE	Tris EDTA
TEMED	N,N,N',N'-tetra methylethylenediamine
TLC	thin layer chromoatography
ts	temperature sensitive
TB	transfer buffer
UV	ultraviolet
UWGCG	University of Wisconsin GCG sequence analysis program
v/v	volume/volume
w/v	weight/volume
X-gal	5-bromo-4-chloro-3-indolylβ-D-galactopyranoside
YE	yeast extract medium

Abstract

The methylation of DNA is a widespread phenomenon found in organisms ranging from bacteria to mammals. Methylation of cytosine at the 5-position is the most common form of this modification and is catalyzed by a conserved family of enzymes. In bacteria, methylation of DNA forms part of the restriction-modification system. The role of DNA methylation in eukaryotes is less clearly defined but it has been implicated in processes such as the control of gene expression and the organization of chromatin structure. Progress in the understanding of DNA methylation could be greatly enhanced by the opportunity to study this phenomenon in genetically tractable organisms such as yeasts and *Drosophila*. However, to date, 5-methylcytosine has not been detected in the DNA of these organisms.

The fission yeast gene *cnd1* + (completion of nuclear division) was cloned by complementation of a temperature-sensitive mutation (R. Bartlett, PhD thesis, Oxford University, 1991) and was found to encode a protein with striking homology to cytosine-specific DNA methyltransferase enzymes (m^5C-MTases). This finding suggested that it might now be possible to study methylation in yeast. Also cloned at this time was an extragenic suppressor of the *cnd1-1* mutant.

In the current work, a more detailed methylation analysis of the fission yeast genome has been carried out. However, it has still not been possible to detect 5-methylcytosine in fission yeast DNA. The extragenic suppressor of the *cnd1-1* mutant has been sequenced and found to encode a small polypeptide with no homology to known proteins. The m^5C-MTase homologue was histidine-tagged, over-expressed in *E. coli*

and purified over a nickel agarose column. The protein did not appear to possess methyltransferase activity when tested with a variety of substrates *in vitro*. However, the lack of activity may be a consequence of the low resolution of the tritium transfer method used in these assays. During analysis of the methyltransferase cDNA, it was discovered that this gene did not rescue the *cnd1-1* mutation. A number of experiments were performed to resolve this anomaly leading to the conclusion that the m^5C-MTase homologue was not encoded by the *cnd1*$^+$ gene. Upon re-examination of the original *cnd1*$^+$ genomic clone, a further gene was found that when transformed alone, was able to rescue the *cnd1-1* mutant. This gene was found to encode the pat1 protein kinase. As the m^5C-MTase homologue was no longer associated with the *cnd1-1* phenotype, it was renamed *pmt1*$^+$ (pombe methyltransferase). Haploid cells lacking the *pmt1*$^+$ gene are viable, indicating that under normal growth conditions *pmt1*$^+$ is not an essential gene. One possible explanation for these results is that *pmt1*$^+$ is a redundant gene, a hypothesis consistent with the lack of detectable methylation in this organism. However, most methods used to detect methylation are relatively insensitive. Northern analysis has shown that the *pmt1*$^+$ gene is still expressed, a finding which would seem to argue against the *pmt1*$^+$ gene function being redundant. Thus the biological significance of the *pmt1*$^+$ gene remains unresolved.

Chapter 1 Introduction

1.1 Characteristics of DNA methylation

i Distribution of DNA methylation

The methylation of DNA is a widespread phenomenon found in organisms ranging from bacteria to humans. It is the major covalent modification found in DNA and can be seen as a way of increasing the information content without altering the nucleotide sequence. In other words, it is an epigenetic modification. Methylation of cytosine at the 5-position (m^5C) is the most common form of this modification and the one that is considered predominantly below. However, N^6-methyladenine and N^4-methylcytosine can also be found, particularly in prokaryotic DNA.

All these modifications occur by the transfer of a methyl group from S-adenosyl-L-methionine (AdoMet) to the relevant base in a reaction catalyzed by a specific DNA methyltransferase. Methyl groups may be added to cytosine residues by either a maintenance methyltransferase or a *de novo* methyltransferase. During replication of methylated DNA, the newly synthesized strand is initially devoid of methyl groups and the resulting hemi-methylated DNA duplex is the preferred substrate of a maintenance methyltransferase. When a previously unmethylated cytosine residue becomes modified, this is known as *de novo* methylation.

Methylation of cytosine at the 5-position is associated with a variety of biological functions. In prokaryotes, methylation distinguishes host DNA from exogenous DNA in the restriction-modification systems that protect bacteria from infection by foreign DNA. Adenine methylation occurs in

some bacterial restriction modification systems and it also plays a role in controlling the initiation of DNA replication. The role of cytosine methylation in eukaryotes is less clearly understood. In the filamentous fungus *Neurospora crassa*, cytosine methylation is associated with the inactivation of DNA duplications in the genome. DNA methylation in higher eukaryotes is implicated in the regulation of gene expression and the organization of chromatin structure.

In higher eukaryotes, m^5C occurs predominantly in the symmetrical dinucleotide CpG and the methyl group is normally present on both strands. Many plant genomes contain even higher levels of m^5C than vertebrates and in addition to the m^5CpG dinucleotides, methylation is also found in CNG triplet sequences where N is any nucleotide. The confinement of m^5C to symmetrical sequences such as CpG and CNG is thought to ensure that patterns of methylation are faithfully transmitted to daughter cells by the activity of a methyltransferase which has a preference for hemi-methylated DNA (Bird, 1978).

ii Effects of methylation on DNA structure

Base modifications can alter the stability of the DNA double helix. Methylation of cytosine stabilizes DNA duplexes whereas methylation of adenine destabilizes the double helix. These changes in stability have been elegantly demonstrated by denaturing gradient gel electrophoresis (Collins and Myers, 1982). In addition, the methylation of cytosine will increase the probability that a region of DNA will assume a Z configuration (Behe and Felsenfeld, 1981). Although the occurrence of Z DNA has been demonstrated *in vitro*, evidence for its existence *in vivo* is still controversial (Jaworski *et al.*, 1988). While such changes in DNA structure may be

functionally significant, the major of consequence of methylation is that DNA-protein interactions will be altered and it is through this latter mechanism that the effects of DNA methylation are thought to be mediated.

1.2 Detection of 5-methylcytosine

In order to study the functions of cytosine methylation, precise and accurate methods are needed to detect individual m^5C residues. A variety of chemical methods exist for assessing the overall m^5C content in DNA. However, to correlate the presence of methylation with specific biological events, techniques are required whereby the methylation status of an individual gene or DNA sequence can be determined. Such techniques usually depend on the differential cleavage of DNA by enzymatic or chemical means. A recently developed method, relies on the chemical conversion of cytosine to uracil by a process which has no effect on m^5C.

i Quantification of overall levels of 5-methylcytosine

To quantify overall levels of m^5C, DNA is degraded into its constituent nucleotides which are then resolved by chromatography. Sensitive methods for quantitative measurements of m^5C in DNA include, high pressure liquid chromatography, (Singer *et al.*, 1977; Kuo *et al.*, 1980), mass spectrometry (Singer *et al.*, 1979), gas chromatography (Razin and Sedat, 1977), thin layer chromatography (Reddy *et al.*, 1981) and stable isotope dilution gas chromatography-mass spectrometry (Crain and McCloskey, 1983). The latter method is probably the most sensitive and is able to detect m^5C at levels of approximately 0.075% of the total cytosine. These techniques have been used to quantify the overall levels of

methylation in the genomes of various organisms. Nevertheless, in order to determine the methylation status of a particular gene, more precise methods are required.

ii Methylation-sensitive restriction enzymes

Perhaps the most common method for analyzing DNA methylation is the use of restriction enzymes which are sensitive to the presence of m^5C in specific sequences (Bird and Southern, 1978; Waalwijk and Flavell, 1978). The DNA is digested with such an enzyme or preferably an isoschizomeric pair of enzymes which have the same recognition sequence but different sensitivities to the presence of m^5C. For example, *HpaII* and *MspI* both recognize the sequence CCGG but only the latter enzyme will cleave the DNA if the internal cytosine is methylated. After digestion, the DNA is resolved by gel electrophoresis and the position of the relevant fragments determined by hybridization using specific probes. The length and intensity of the DNA fragments resulting from the various enzymes will provide information as to the methylation status of the particular restriction sites.

An extension of this assay has been developed using the Polymerase chain reaction (PCR) to provide a more quantitative means of assessing DNA methylation (Singer-Sam *et al.*, 1990). Once again, DNA is cut with an isoschizomeric pair of restriction enzymes. The DNA is amplified by PCR using primers flanking the restriction site of interest so that only intact and presumably methylated DNA will generate the expected fragment. The resulting DNA products are then resolved by gel electrophoresis and quantified by densitometry.

Despite the wealth of information that can be obtained from the use of methylation-sensitive restriction enzymes, there are disadvantages to this

method. Most isoschizomeric pairs recognize cytosine as part of a symmetrical sequence, for example CCGG or CCWGG (where W is A or T) and thus only a small proportion of possible CpG or CNG sequences will be analyzed. In addition, the detection of methylation is based on a negative observation, namely the lack of cleavage and this has the potential to give ambiguous results.

iii Selective chemical cleavage

There are a number of sequencing methods for the detection of m^5C and these are based on the Maxam and Gilbert chemical cleavage methodology (Maxam and Gilbert, 1980). With this technique, hydrazine modifies thymine and cytosine residues but does not react with m^5C and consequently, this latter base is protected from subsequent piperidine cleavage. This leads to a gap in the resulting sequencing ladder which can be ascribed to the presence of m^5C. Various alterations have been made to improve this method including an indirect end-labelling protocol that can be applied to genomic DNA (Church and Gilbert, 1984), the incorporation of a PCR amplification step (Saluz and Jost, 1986), and a ligation-mediated PCR (LMPCR) technique which lowers the amount of genomic DNA required (Mueller and Wold, 1989; Pfeifer *et al.*, 1989).

Many of these techniques, however, are lengthy and require relatively large amounts of genomic DNA. In addition, as with the use of isoschizomer analysis, the recognition of m^5C is by a negative observation, in this case a gap in the sequencing ladder. A positive signal can be obtained with the permanganate cleavage method (Fritzsche *et al.*, 1987) although this method is also susceptible to many problems and has not been used to analyze genomic DNA.

iv Bisulphite sequencing method

This recently developed bisulphite sequencing technique is probably the most precise and sensitive way to determine the methylation status of a particular DNA sequence (Frommer *et al.*, 1992). Treatment of single stranded DNA with bisulphite will result in the deamination of cytosine to uracil whereas m^5C will remain unconverted. PCR is then used to amplify the DNA with the result that uracil will be amplified as thymine while m^5C will still be recognized as cytosine. As the conversion of cytosines to uracils creates non-complementary base pairs (uracil opposite guanines), DNA can be amplified using two separate pairs of primers which are specific for each strand. The PCR product can be sequenced directly in which case the average methylation of a particular cytosine is assessed. Alternatively, PCR products can be cloned and then sequenced indicating the methylation status of an individual cytosine.

This technique has several advantages over the previously described methods. Firstly, it results in a positive identification for m^5C (in the form of a band on a sequencing gel). Furthermore, as already mentioned, this signal can represent an individual cytosine or an average for several clones at that position. Finally and perhaps the most important aspect of this method is that only small amounts of DNA are required, for example, the DNA from less than 100 mammalian cells is sufficient for the methylation analysis of a single copy gene.

Drawbacks to this technique include the variable nature of the bisulphite reagent and the stringent protocols required for its removal after the conversion reaction. A further critical parameter is the design of the primers as this can greatly influence the effectiveness of the whole method (Grigg and Clark, 1994).

1.3 The cytosine-5 DNA methyltransferase family

The cytosine-5 methyltransferases (m^5C-MTases) are a conserved family of enzymes that transfer methyl groups from S-Adenosyl-methionine (AdoMet) onto the 5-position of cytosine in DNA. Most of the enzymes described to date are from prokaryotic species, although a few eukaryotic examples have now been identified. This family shares a common architecture consisting of ten conserved amino acid motifs which alternate with non-conserved regions (Posfai *et al.*, 1989). Six of these motifs are considered to be highly conserved and the remaining four motifs are considered moderately conserved. The motifs are between three and twenty amino acids long and are present in a conserved order.

Motif I is similar to a weakly conserved sequence, present in other methyltransferases that use AdoMet, and is thought to be required for interactions with this cofactor (Klimasauskas *et al.*, 1989; Ingrosso *et al.*, 1989). Motif IV contains an invariant cysteine residue that is essential for catalysis. Wu and Santi (1985) initially proposed that the mechanism of m^5C-MTases is similar to that employed by thymidylate synthase (Santi and Danenberg, 1984) and tRNA uracil 5-MTase (Santi and Hardy, 1987). Both these enzymes transfer a methyl group to the C5 position of a pyrimidine by a mechanism that involves a cysteine residue. The thiol group of this cysteine serves as a nucleophile that attacks the 6-carbon of cytosine resulting in a DNA-protein intermediate. The addition of the nucleophile activates the 5-carbon thus enabling the methyl group to be transferred from AdoMet, accompanied by release of S-adenosyl-L-homocysteine. After transfer of the methyl group, the proton at the 5-position is abstracted by a basic residue on the enzyme.

Evidence that the conserved cysteine in region IV is the catalytic site has come from a variety of experiments on prokaryotic enzymes. For example, site directed mutagenesis experiments resulting in the substitution of cysteine by a number of other amino acids, leads to a loss of m^5C-MTase activity in *EcoRII* (Wyszynski *et al.*, 1992), *HaeIII* (Chen *et al.*, 1993), Dcm (Hanck *et al.*, 1993) and *HhaI* (Mi and Roberts, 1993). The substrate 5-fluorodeoxycytosine has been useful in studying m^5C-MTases as it forms a covalent bond to the enzyme during the methyl transfer reaction, resulting in the trapping of reaction intermediates. This bond is formed between the cysteine residue and the carbon-6 of 5-fluorodeoxycytosine (Osterman *et al.*, 1988; Chen *et al.*, 1991; Friedman and Ansari; 1992; Smith *et al.*, 1992).

The DNA recognition domain lies between motifs VIII and IX and is between 80 and 300 amino acids long. It is responsible for the recognition of a specific DNA sequence containing the target cytosine (Wilke *et al.*, 1988; Klimasauskas *et al.*, 1991). For prokaryotic enzymes, this sequence is usually between four and eight base pairs long. Interchanging the recognition domains of m^5C-MTase enzymes shows that these regions determine not only the sequence specificity, but also the choice of base within that sequence (Klimasauskas *et al.*, 1991, Mi and Roberts, 1992).The enzymes *BsuRI*, *BspRI* and *HaeIII* all recognize the sequence GGCC and the amino acid sequences of the variable regions in these proteins are similar (Posfai *et al.*, 1989). However, the variable regions of other enzymes that recognize identical target sequences such as *MspI* and *HpaII* are not similar and therefore it is not possible to predict the target sequence of a m^5C-MTase from its variable domain sequence.

Remarkable progress in the understanding of the structure of m^5C-MTases and their mechanism of action was recently made by the determination of the crystal structure of the *HhaI* m^5C-MTase. Initial

experiments characterized the crystal structure of *HhaI* complexed with the cofactor AdoMet (Cheng *et al.*, 1993). A second more intriguing study resolved the crystal structure of a chemically-trapped covalent reaction intermediate between *HhaI*, S-adenosyl-L-homocysteine and a duplex DNA oligonucleotide containing methylated 5-fluorocytosine as its target base (Klimasauskas *et al.*, 1994). From the X-ray crystallography data, it appears that the protein is folded into two domains with the conserved motifs clustered in the core of the larger domain. The variable target recognition region is contained within the smaller domain and the substrate DNA is located in a cleft between the two domains. Interestingly, the target cytosine residue is swung completely out of the DNA helix so that it becomes positioned in the active site which itself undergoes a large conformational change during catalysis.

The first eukaryotic m^5C-MTase to be identified and extensively characterized, is from the mouse (Bestor *et al.*, 1988). This enzyme methylates cytosine in CpG dinucleotides. Subsequently, m^5C-MTase genes were cloned from human (Yen *et al.*, 1992) and *Arabidopsis thaliana* (Finnegan and Dennis, 1993). All three eukaryotic enzymes possess a carboxy-terminal domain that is closely related to the bacterial enzymes. In addition, the eukaryotic enzymes possess an amino-terminal extension of approximately 1000 amino acids which shows no homology to known proteins. The amino-terminal domain is thought to regulate the activity of the m^5C-MTase and confer upon it the specificity for hemi-methylated DNA (Bestor, 1992). Within the N-terminal region of the mouse enzyme is a novel cysteine-rich region that has been shown to bind zinc ions. This zinc binding motif is also present in the human enzyme but not in the *Arabidopsis* m^5C-MTase.

The preference of the mouse enzyme for hemi-methylated DNA raises the question of how the *de novo* methylation, that occurs in germ cells

and early embryos, is catalyzed. One possibility is that another as yet unidentified m^5C-MTase exists during these developmental stages. Alternatively, a different form of the existing m^5C-MTase might carry out such *de novo* methylation.

Plant DNA contains methylation at CNG sequences (where N is any nucleotide) as well as in CpG dinucleotides. It is not known whether these different methylation events are catalyzed by the same m^5C-MTase or whether separate enzymatic activities are required. Certain prokaryotic enzymes are multispecific in that they recognize and methylate more than one target sequence. It may also be that the eukaryotic enzymes are capable of methylating cytosine residues in more than one sequence context. To date, there have been no reports of enzymatic studies using purified recombinant eukaryotic enzymes. The development of such systems would greatly assist with the functional studies of eukaryotic m^5C-MTase enzymes.

1.4 Mutability of 5-methylcytosine

The presence of m^5C in DNA could be considered surprising given its ability to act as a mutagen. Deamination of cytosine generates uracil, a base normally found in RNA and therefore recognized as being foreign. It is removed from the DNA by the enzyme uracil DNA glycosylase and subsequently repaired by the normal repair pathway. In contrast, deamination of m^5C results in the formation of thymidine, a normal component of DNA which is not recognized by such a system. Both *E. coli* and mammalian cells possess a system capable of recognizing T/G mismatches and repairing them in favour of G (Jones *et al.*, 1987; Brown and Jiricny, 1987). However, even though such repair mechanisms exist, the deamination of m^5C often results in a CpG to ApT transition. For example,

three sites in the *E. coli lac I* gene were found to be hotspots for m^5C to T transitions. These sites were not mutagenic in non-methylating strains indicating that the presence of m^5C was responsible for the transition (Coulondre *et al.*, 1978).

The CpG dinucleotide is present at approximately 20% of its expected frequency in vertebrate genomes (Bird, 1980). It is thought that this deficiency is caused by the high mutation rate of m^5CpG to TpG and CpA by deamination. A comparison between the human pseudo α-globin gene sequence and the functional gene showed that the majority of CpGs in the latter had mutated to CpA or TpG in the pseudo-gene (Bird *et al.*, 1987). Sved and Bird (1990), have proposed that the levels of CpG represent an equilibrium between rate of creation of new CpGs by point mutation and the loss of CpG by deamination of m^5CpG.

The finding that m^5C is an effective mutagen would seem to suggest that its maintenance in the genomes of many species is due to selective pressures. Although in many cases the exact role of DNA methylation is unclear, increasing evidence has accumulated over recent years to support the view that DNA methylation does have specific biological functions.

1.5 DNA methylation in prokaryotes

The role of DNA methylation in prokaryotes has been extensively characterized. For example, in *E. coli*, *dam* methylation is involved in the control of DNA mismatch repair (Messer and Noyer-Weidner, 1988; Modrich, 1991). The product of the *dam* gene is a DNA methyltransferase which modifies adenine in the sequence GATC. After the replication fork has passed through a region of DNA, there is a delay before the newly synthesized strand is methylated. During this time lag, errors incorporated

into this strand can be corrected by the mismatch repair system. The lack of methylation acts as a signal to direct the repair machinery to the newly synthesized strand with the result that errors are corrected before the next round of replication thereby preventing their inheritance.

dam$^-$ strains are viable but display a "mutator" phenotype whereby many mutations are incorporated into the DNA. The same effect is apparent in strains over-expressing the *dam* gene. In this case, overproduction of the methyltransferase results in the full modification of DNA before the repair system can act. As well as this role in mismatch repair, evidence is accumulating that Dam methylation in *E. coli* is also involved in the regulation of processes such as DNA replication, chromosome segregation and gene expression (Messer and Noyer-Weidner, 1988).

Despite these diverse functions, the most important role of DNA methylation in bacteria appears to be in conferring immunity to bacteriophage infection. Methylation is part of the restriction-modification (R-M) system that is used to distinguish between host and foreign DNA (reviewed in Wilson and Murray, 1991). These systems are found mainly in bacteria, although they are encoded by some bacterial viruses (Humbelin *et al.*, 1988) and also by certain *Chlorella* viruses (Van Etten *et al.*, 1989).

There are three main types of R-M systems which are categorized according to enzyme composition, symmetry of recognition sequence, position of cleavage and cofactor requirement. Type II is the simplest system and the one with the most members. The genes encoding the restriction and modification components of this system are normally closely linked in the genome but unlike other R-M systems, each component functions as a separate enzyme.

The type II modification component is a DNA methyltransferase which modifies cytosine or adenine residues in a specific symmetrical sequence which is usually between four and eight base pairs long. This

sequence is also recognized and cleaved by the restriction component, an endonuclease but this enzyme is unable to cleave the DNA when it is methylated. Bacterial DNA is usually fully modified by the modification component, however exogenous DNA that enters the cell, for example by bacteriophage infection, is non-methylated and therefore cleaved which renders it non-infectious.

1.6 DNA methylation in eukaryotes

i Patterns of methylation

The patterns and distributions of DNA methylation in eukaryotes are very different from those found in bacterial species. In prokaryotes, all sequences that are recognized by a particular DNA methyltransferase are fully methylated. Amongst eukaryotes, only m^5C is found in appreciable amounts and the proportion of cytosine which is modified varies considerably between species. In eukaryotic species that possess methylated DNA, the modified regions are often interspersed with non-methylated DNA. The distribution and amount of methylation found in the DNA of eukaryotic species can be divided into three broad categories (Bird, 1980).

The highest levels of methylation found within the animal kingdom are present in the genomes of vertebrates such as mouse and man where between 3-8% of cytosines in DNA are methylated (Shapiro, 1975).

The second category of eukaryotic DNA methylation includes organisms whose genomes possess intermediate levels of methylation, for example sea urchin and slime moulds. The final class includes species in which m^5C is either absent from or below the level of detection in the DNA. Eukaryotes that fall into this latter group include *Drosophila melanogaster*

(Urieli-Shoval *et al.*, 1982), *Caoenorhabditis elegans* (Simpson *et al.*, 1986) and yeasts (Proffit *et al.*, 1984; Antequera *et al.*, 1984).

ii DNA methylation in vertebrates

Methylation is spread throughout the vertebrate genome and can be found in both transcribed and non-transcribed regions of DNA. Approximately 70% of cytosines in CpG dinucleotides are methylated but a small distinct fraction of the genome (1-2%) can be identified which contains non-methylated CpGs (Cooper *et al.*, 1983). These non-methylated regions are between 0.5 and 2kb long and possess a G + C content of 60% which is significantly higher than that found in bulk DNA. In addition, the CpG dinucleotide is found at the expected frequency (Bird, 1986). These methylation-free sequences are known as "CpG islands" and are normally associated with the 5' domains of promoter regions of all constitutively expressed (housekeeping) genes and a small proportion of tissue specific genes representing in total, approximately 56% of all human genes (Gadriner-garden and Frommer, 1987; Antequera and Bird, 1993).

CpG islands remain free of methylation in all cell types though it is not understood how they escape this modification. It may be that their ability to bind transcription factors such as Sp1 prevents the m^5C-MTase gaining access to these sequences (Macleod *et al.*, 1994). The only islands that are methylated include those on the inactive X chromosome of female mammals and those associated with genes that are not essential for growth in tissue culture cell lines (Antequera *et al.*, 1990). In addition, islands associated with some genes subject to genomic imprinting (described below) are also methylated.

Evidence has been accumulating for some years suggesting that the primary function of DNA methylation is to repress gene expression. The proliferation of many eukaryotic viruses is prevented when they become methylated. For example, the methylation of adenovirus 2 correlates with its transcriptional silencing (Sutter and Doerfler, 1980) and the methylation of other viral promoters has been found to cause their inactivation (Vardimon *et al.*, 1982). Injection of retroviral proviruses into mouse embryos also results in transcriptional inactivation by methylation (Jaenisch, 1982). This repression is maintained throughout development by the continued methylation of these sequences. The *in vitro* methylation of the CpG islands in a number of transcription units has been found to inhibit gene expression following transfection into fibroblast cells (Razin and Cedar, 1991) or after incubation in cell extracts (Vardimon *et al.*, 1982). In addition, many endogenous genes whose expression is normally repressed can be activated by treatment with 5-azacytidine, an analog of cytosine whose incorporation into DNA results in hypomethylation (Jones, 1984).

In some cases, however, it appears that gene inactivation precedes methylation *in vivo* suggesting that DNA methylation is a consequence rather than a cause of repression, merely serving to reinforce the inactive state of a particular sequence (Gautsch and Wilson, 1983). Consistent with this theory is the activation of certain methylated genes despite the lack of change in their methylation pattern (Kelley *et al.*, 1988). For many years the exact role of DNA methylation in vertebrates has remained a controversial issue, however its importance has been clearly demonstrated by the finding that it is essential for normal embryonic development in mice (Li *et al.*, 1992). Embryonic stem (ES) cells carrying disrupted copies of the DNA methyltransferase gene were used to generate mouse embryos which were found to possess only one-third of wild type levels of genomic m^5C.

Whether the remaining methylation is the result of the activity of a different DNA methyltransferase or merely due to incomplete disruption of methyltransferase activity is not clear. The embryos fail to develop past mid-gestation indicating that DNA methylation is essential for normal development. Interestingly, undifferentiated ES cells with both copies of the DNA methyltransferase gene disrupted are viable, suggesting that high levels of methylation are not required for cell viability.

Mature germ cells of mice are known to contain high levels of DNA methylation which are thought to greatly decrease upon fertilization. Subsequent *de novo* methylation and demethylation steps during embryogenesis are thought to contribute to the establishment of the patterns seen in the adult (Kafri *et al.*, 1992; Monk *et al.*, 1987). If such changes in the distribution of DNA methylation contribute to the regulation of gene expression in embryogenesis, the disruption of the DNA methyltransferase and the subsequent reduction in methylation would be expected to severely affect the development process. However, the correlation of methylation with tissue specific gene expression during embryogenesis remains to be established.

There are other areas of vertebrate development where DNA methylation has been implicated in the regulation of gene expression. Firstly, evidence has accumulated supporting the view that DNA methylation is involved in the process of X chromosome inactivation (reviewed in Riggs and Pfeifer, 1992). This process occurs in female placental mammals and involves the inactivation of most genes on one of the two X chromosomes. The inactivation is correlated with the methylation of the CpG islands of these genes.

Secondly, genomic imprinting is another case where CpG islands are thought to be methylated during development. Genomic imprinting is the

phenomenon whereby alleles of different parental origin are differentially expressed (reviewed in Monk, 1987). This process requires that the DNA is marked in a heritable yet reversible manner. Methylation has the required epigenetic characteristics to be the imprinting signal. However, while a number of studies have implicated methylation in the control of imprinting, its exact role in this process is not clear. It may be that methylation is not the actual imprint signal itself, but that it contributes to the regulation of imprinted genes (Sasaki *et al.*, 1992).

By what mechanism does DNA methylation repress transcription? There seem to be two possibilities. Firstly, methylation could act directly by preventing the interactions between transcription factors and their binding sites. A number of factors have been identified which bind to sites containing CpG dinucleotides and some of these factors do not bind when the cytosine residue is methylated (Kovesdi *et al.*, 1987; Watt and Molloy, 1988; Iguchi-Ariga and Schaffner, 1989) resulting in an inhibition of transcription. An alternative, more indirect mechanism for the action of methylation is through proteins which bind to methylated DNA. One such methyl-CpG binding protein (MeCP1) has been characterized by Meehan *et al.*, (1989, 1992). This protein binds specifically to DNA sequences that contain multiple (at least 10) methylated CpG sites. Beyond this requirement, MeCP1 has no apparent sequence specificity, a finding which suggests that it could mediate the effects of DNA methylation in a variety of sequence contexts. In support of this theory, MeCP1 has been shown to repress transcription *in vitro*. Interestingly, in ES cells where MeCP1 is present at very low levels, transcription cannot be repressed by methylation (Boyes and Bird, 1991).

A second methyl-CpG binding protein, MeCP2, has also been identified (Lewis *et al.*, 1992). This protein is able to bind to DNA that

contains a single methyl-CpG pair and is far more abundant than MeCP1. *In situ* immunofluorescence in mouse cells has shown that although MeCP2 is present throughout the chromosome arm, it is concentrated in the pericentromeric heterochromatin. This latter region contains half of all genomic m^5C and is also the location of the mouse major satellite DNA. This localization could indicate that MeCP2 contributes to the formation of compact chromatin structures although an additional role in transcriptional repression cannot be excluded at this point (Meehan *et al.*, 1992).

There is increasing evidence that DNA methylation is a determinant of higher order chromatin structure in mammals. In this way, methylation could influence transcriptional repression by regulating chromatin structure such that accessibility to certain genes is reduced. Following DNA mediated gene transfer into animal cells in culture, unmethylated exogenous DNA integrates into the genome in a DNase I sensitive conformation. By contrast, an inactive chromatin structure that is resistant to nuclease digestion is adopted by exogenous methylated DNA (Keshet *et al.*, 1986).
In addition, when naked DNA is injected into cells, transcription occurs equally from methylated and unmethylated templates. However, as soon as the DNA becomes packaged, transcription from the methylated template is inhibited. Furthermore, if the same experiment is performed using preformed chromatin templates, the inhibitory effect of methylation is seen at once (Buschhausen *et al.*, 1987).

Vertebrate methylation has long been thought to occur only in the symmetrical CpG dinucleotide. However, recent experiments on mammalian origins of replication have challenged this view. Densely methylated regions were observed in two out of four chromosomal replication origins from Chinese hamster cells using the bisulphite genomic sequencing method (Tasheva and Roufa, 1994). The methylation is

correlated with DNA replication as when cell growth and replication are arrested, the majority of cytosines are unmodified. Interestingly, the dense methylation of cytosine occurs at CpA, CpT and CpC dinucleotides as well as CpG. What could be the role of this methylation in DNA replication ? One possibility is that methylation mediates association between the origins and the nuclear envelope (adenine methylation is thought to influence a similar process in bacteria). Alternatively, the methylation may determine whether a particular origin is to be activated at all since far more origins exist than can be used in any single round of DNA synthesis. Finally, the methylation could mark a particular origin to indicate to the replication machinery that it has fired and should not be used again in that particular round of S phase. However, as not all origins of replication possess methylated regions, the exact significance of the methylation is unclear. As described below, non-symmetrical methylation has been reported in other eukaryotic species and it will be interesting to learn how widespread this type of methylation is and what its functions are in eukaryotic biological processes.

iii DNA methylation in plants

The genomes of many plant species have been found to contain substantial amounts of m^5C and in some cases as much as 30% of the total cytosine is modified (Shapiro, 1975). This amount is significantly higher than the levels of m^5C found in vertebrates where 3-8% of cytosines are methylated and it is thought that this difference is attributable to two main factors. Firstly, DNA methylation is generally confined to cytosines in CpG dinucleotides whereas in plants methylation can be found in cytosines that are part of both CpG dinucleotides and CNG triplets (where N is any

nucleotide, Gruenbaum *et al.*, 1981). In addition, the frequency of the CpG dinucleotide in plant DNA is far higher than that found in animals.

The genomes of a number of plant species contain a non-methylated fraction that is rich in CpG dinucleotides (Antequera and Bird, 1988). These regions were found to be associated with the 5' ends of certain genes in a similar manner to the vertebrate genomes discussed above. The possible role of methylation in the regulation of plant gene expression has not been explored in the same depth and detail as the studies with vertebrate genes and it remains to be seen whether DNA methylation has an analogous role in the plant kingdom or whether the higher levels of methylation are associated with additional functions.

Recently, the cloning of the first DNA methyltransferase gene from a plant species, *Arabidopsis thaliana*, was reported (Finnegan and Dennis, 1993). This DNA methyltransferase is very similar in overall structure to the mouse and human enzymes and further analysis of this enzyme should help to elucidate the role of DNA methylation in plants. Most of the m^5C in the *Arabidopsis* genome is found in the highly repeated DNA fraction (Pruitt and Meyerowitz, 1986). This includes the 180bp centromere repeat family and the ribosomal RNA genes. Other sequences known to be methylated are the moderately reiterated and single-copy telomere-associated sequences.

Arabidopsis mutants with reduced levels of DNA methylation, *ddm1* and *ddm2*, were isolated by screening mutagenized populations of plants (Vongs *et al.*, 1993). Both mutants were found to develop normally and displayed no striking morphological phenotypes. However, upon pollination of the mutant strains, a segregation distortion phenotype was discovered whereby progeny containing hypomethylated genomes were over-represented in F2 generations produced by self pollination of

ddm1-1 /+ and *ddm1-2* /+ heterozygotes. It is not clear how a reduction in DNA methylation could result in such a phenotype but suggests a possible role for m^5C in processes such as meiosis or gametogenesis. Although the reduction in methylation does not affect development, these mutants still contain 25-30% of wild type levels of m^5C and therefore it may be that some of this residual methylation is essential for normal growth.

One role of DNA methylation in plants seems to be in controlling the activity of transposable elements. For both the maize Activator (Ac) and Suppressor-mutator (Spm) transposable elements, methylation of a region upstream and downstream of the site of initiation of transcription results in a loss both of mRNA encoding for the transposase and the ability to excise (Dennis and Brettell, 1990). Furthermore, an inactivated methylated Ac derivative was activated by cultivation of embryos in tissue culture. Plants that were heterozygous for the inactivated allele were self-fertilized and after pollination, immature embryos were dissected out and placed into tissue culture. Plants were re-generated from tissue culture and found to contain an active element. The 5' end of the corresponding transposase gene was found to be demethylated. The frequency of reactivation of inactive elements is greatly increased by the introduction of an active element into the genome (Dennis and Brettell, 1990). One possible explanation for this finding, is that the transposase made by the active element binds to the inactive element and prevents methylation of the critical regions.

Recent experiments have provided evidence for cytosine methylation of non-symmetrical sequences in a transgenic line of *Petunia hybrida* (Meyer *et al.*, 1994). Using the bisulphite genomic sequencing technique, the cytosine methylation patterns within a maize transgene, that had integrated into the *Petunia* genome, were analyzed. The sequence of both a hypermethylated and a hypomethylated allele were examined and both were found to contain

non-symmetrical as well as symmetrical m^5C residues within the promoter of the transgene. Non-symmetrical m^5C residues were present at frequencies of 5.9 and 31.9% in the hypomethylated and hypermethylated states respectively. To date, it is not clear whether non-symmetrical methylation is confined to transgenes or whether it is also present in endogenous plant DNA. It will be interesting to determine whether the symmetrical and non-symmetrical methylation patterns are established by the same methyltransferase. Low stringency Southern blot analysis has revealed that a small multigene family with homology to m^5C-MTases exists in *Arabidopsis*. Furthermore, a new putative methyltransferase gene from this species was recently identified (Scheidt *et al.*, 1994). Intriguingly, this gene encodes a protein that resembles a bacterial rather than a eukaryotic enzyme as it does not possess the N-terminal extension. Moreover, this new m^5C-MTase gene is expressed in the callus and during germination but not in one month old plants or in leaves, suggesting that putative m^5C-MTase gene might be responsible for *de novo* methylation in the early stages of *Arabidopsis* development.

iv DNA methylation in invertebrates

The genomic DNA of eukaryotic species which possess intermediate levels of methylation appears to be organized into methylated and non-methylated compartments. This finding is best illustrated in species such as the sea urchin and the slime mould *Physarum polycephalum*.

Several sea urchin gene families are found in the non-methylated fraction regardless of their expression, whereas repetitive sequences resembling satellite DNA, are only found in the methylated fraction (Bird *et al.*, 1979). Approximately 20% of the *Physarum* genome is

methylated (Rothnie *et al.*, 1991). This fraction consists almost exclusively of copies of a retrotransposon-like sequence Tp1. The presence of this element is thought to be tolerated if transposition is repressed by methylation.

v DNA methylation in fungi

A greater understanding of DNA methylation in higher organisms could be obtained by studying its role in genetically tractable eukaryotes such as yeasts and *Drosophila*. However, as mentioned above, m^5C has not been detected in the DNA of *Schizosaccharomyces pombe* (Antequera *et al.*, 1984), *Saccharomyces cerevisiae* (Proffit *et al.*, 1984) or *Drosophila melanogaster* (Urieli-Shoval *et al.*, 1982).

While m^5C has not been detected in the DNA of yeasts, it has been found in other species of fungi. For example, *Neurospora crassa* and *Phycomyces blakesleeanus* have 1.5% and 0.5% of total cytosine modified respectively (Russell *et al.*, 1985; Antequera *et al.*, 1984). There are reports of methylation in species representing each division of the fungal kingdom (Antequera *et al.*, 1984; Russell *et al.*, 1985; Russell *et al.*, 1987a and Zolan and Pukkila, 1985), but the occurrence of methylated DNA in fungi is somewhat sporadic and there is no obvious evolutionary relationship between the species in which it can be found.

In the filamentous fungus, *Neurospora crassa*, methylation is involved in an interesting premeiotic process that appears to test every sequence in the genome against all others (Selker *et al.*, 1987; Selker and Garret, 1988). This process was identified when it was noticed that although genetic markers, that have been transformed into cells and become integrated into the genome, are expressed normally in vegetative growth, they often become inactivated when strains are crossed with one another. In this

species, nuclear fusion does not occur immediately after fertilization so it would appear that individual nuclei within the dikaryon can detect whether additional copies of a DNA sequence are present. Duplicated sequences are then peppered with GpC to ApT mutations, a process which inactivates both copies permanently, hence the name "Repeat-Induced Point mutation" or RIP (Cambareri *et al.*, 1989). The affected sequences are also left densely methylated and this modification persists in vegetative cultures. Methylation is is somehow induced by the mutations and is not limited to symmetrical sequences such as CpG and CNG, (Selker *et al.*, 1993). A tandem repeat called zeta-eta, consisting of an imperfect duplication of a 5S RNA gene, is highly mutated and methylated and is thought to represent an ancient "Ripping" event. If a copy of the zeta-eta repeat is de-methylated by propagation in *E. coli* and then transformed into vegetatively growing *Neurospora*, the transformed copy becomes re-methylated, even if the parent strain lacks a copy of this region (Selker *et al.*, 1987). This suggests that the methylation is most likely induced by the mutations themselves (Selker, 1990).

The mechanism by which duplicated sequences are detected and inactivated while endogenous repeated sequences such as the rDNA genes are left alone, is not yet understood. A similar process occurs in the related fungus *Ascobolus immersus*, however inactivation in this case is achieved by methylation alone and this is known as "Methylation Induced Premeiotically" or MIP (Goyon and Faugeron, 1989, Goyon *et al.*, 1994). As with RIP, methylation of cytosines during the MIP process is not confined to symmetrical sequences. This methylation is maintained through mitotic and meiotic divisions implying that a novel type of maintenance methyltransferase, which recognizes non-symmetrical cytosine methylation, exists in these cells. In contrast to RIP, if a DNA segment that has undergone

MIP is stripped of its methylation, it remains unmethylated when re-introduced into *Ascobolus*. The fact that inactivation of sequences by MIP is reversible and that no detectable sequence alterations appear to be associated with methylation, strongly implies that methylation alone is responsible for gene inactivation in MIP. A similar process to MIP has recently been identified in the basidiomycete *Coprinus cinereus* although the level of methylation associated with gene inactivation in this fungus appears to be lower and m^5C residues are predominantly found in the CpG dinucleotide (Freedman and Pukkila, 1993).

What are the possible functions of RIP and MIP? It has been suggested that MIP might control meiotic recombination thereby preventing potentially lethal rearrangements of the genome (Faugeron *et al.*, 1990). RIP and MIP could also be ways of inactivating transposons, preventing them from multiplying in the genome. Transposons have been identified in *Neurospora* (Kinsey *et al.*, 1989) and genetic evidence has implied that they also exist in the genome of *Ascobolus* (Decaris *et al.*, 1979; Nicolas *et al.*, 1987).

Mutations affecting the overall levels of methylation in *Neurospora* have been identified by random mutagenesis (Foss *et al.*, 1993). A mutation in one gene, *dim2*, results in the loss of all detectable methylation (the identity of the *dim2* gene product is not known). The *dim2* mutant is viable and the apparent lack of methylation in its DNA suggests that DNA methylation is not essential for viability of *Neurospora*. However, the possibility cannot be ruled out that these cells contain residual levels of methylation, below the limits of detection, which are essential for growth. Abnormal segregation of the methylation levels in crosses led to the finding that *dim2* mutants frequently give rise to progeny containing extra chromosomes or chromosomal parts. This phenotype can also be induced by starving wild type cells of AdoMet, the methyl group donor for DNA

methylation reactions. It would appear, therefore, that DNA methylation is somehow involved in the regulation of chromosome behaviour in *Neurospora*.

Methylated cytosine residues have also been found in the DNA of a number of other species of fungi, however the function of methylation in these cases has not been extensively characterized. Specht *et al.*, (1982) have shown that while most of the genome of the basidiomycete *Schizophyllum commune* is unmethylated, significant methylation of cytosines occurs in the ribosomal RNA genes. Furthermore, these rDNA sequences are more highly methylated in the dikaryon than in the homologue stage. Cytosine methylation has also been found in the spacer regions between rDNA genes in *Neurospora* (Perkins *et al.*, 1986). Another example of methylated DNA in a fungal species is from the basidiomycete *Coprinus cinereus* , where the DNA sequences around the centromere regions were found to contain m^5C (Zolan and Pukkila, 1986).

Different levels of methylation have been reported in various developmental stages of certain fungi. The yeast form of *Candida albicans* was found to have over twice the amount of m^5C as the mycelial form (Russell *et al.*, 1987a). In *Neurospora*, the level of m^5C in the asexual spores was found to be nearly twice that of the mycelial form (Russell *et al.*, 1987b) and the spores of *Phycomyces blakesleeanus* contain six times the amount that is found in the mycelial cells (Antequera *et al.*, 1985). Higher levels of m^5C were found in the DNA from the dormant sclerotia of *P. omnivorum* than in the DNA from the metabolically active mycelia (Jupe *et al.*, 1986). It remains to be established whether there is a link between these differences in methylation and the regulation of specific genes involved in the developmental changes. It is an attractive hypothesis though, that the higher levels of methylation seen in the more dormant stages of the life cycles in

these fungal species, might be responsible for repressing gene functions that are required for vegetative growth.

As mentioned above, m^5C has not been found in the DNA of the yeasts *S. cerevisiae* or *S. pombe* (Proffit *et al.*, 1984; Antequera *et al.*, 1984). Efforts to detect methylation in these species used high pressure liquid chromatography for *S. cerevisiae* DNA and isoschizomeric restriction enzyme analysis and thin layer chromatography for *S. pombe*. As described in section 1.2, these methods have a resolution such that m^5C will only be detected if it is present at levels of 0.075-0.1% of the total cytosine. It is therefore possible that small amounts of m^5C may exist in the genomes of these species.

During the course of evolution, DNA methylation appears to have acquired new roles and patterns of distribution from those seen in prokaryotes. While DNA methylation is involved in a simple immune system in bacteria, it appears to play a role in regulating the structure and expression of the genome in higher eukaryotes. One common theme in eukaryotic DNA methylation is the inactivation of extraneous DNA sequences, such as transposable elements and repetitive DNA. In addition, it is likely that by regulating chromatin structure in higher eukaryotes, DNA methylation serves to compartmentalize the genome in order to reduce the amount of DNA that has to be scanned by transcription factors and other DNA binding proteins. The methylation of repetitive sequences at centromeres and telomeres in some organisms may also play a role in regulating chromosome segregation and behaviour.

1.7 The cloning of a fission yeast gene that encodes a m^5C-MTase homologue

The identification of a temperature-sensitive (ts) lethal mutant defective in a putative m^5C-MTase homologue raised the possibility that there is an essential function for m^5C in *S. pombe* (Bartlett, 1991). The ts lethal strain was isolated from a screen for mutants that had undergone re-replication of their DNA. At the restrictive temperature, cells accumulated with two closely opposed nuclei. After a few more hours at 36ºC, some of the cells progressed through mitosis such that there were now four closely opposed nuclei. Due to this failure to complete nuclear division, the mutant was called *cnd1-1*. In addition, *cnd1-1* cells at the permissive temperature were found to be capable of conjugation and sporulation, even when growing on rich media, implying that the mutation was bypassing the requirement for starvation before entering the meiotic pathway.

In order to clone the *cnd1*$^+$ gene, two different genomic libraries were transformed into the *cnd1-1* mutant. One library contained a complete *HindIII* digest of genomic DNA in the vector pDB262 (Wright *et al.*, 1986), whereas the other contained a partial *SauIIIA* digest of *S. pombe* genomic DNA in the vector pYEP13 (Broach *et al.*, 1979). Two different plasmids were isolated (one from each library), that were able to rescue the ts phenotype. The rescuing plasmid from the pDB262 library (C1), was found to contain a single 2.4kb insert. When the C1 plasmid was transformed into the *cnd1-1* mutant, some colonies of uniformly growing cells were seen, suggesting that the plasmid was integrating into the genome and resulting in the stable rescue of the *cnd1-1* mutant. The stably-rescued strain was crossed to a wild type strain. Random spore analysis showed that nearly all of the colonies were able to grow at the restrictive temperature. This tight linkage between

the *cnd1-1* mutation and the rescuing plasmid suggested that the C1 plasmid contained the *cnd1*$^{+}$ gene and that the plasmid had integrated at the site of the *cnd1-1* mutation by homologous recombination. The rescuing plasmid from the pYEP13 library contained a different insert to that found in C1 and was presumed to encode a suppressor of the *cnd1-1* mutant.

The 2.4kb *HindIII* insert from the C1 plasmid was sequenced and found to contain one major open reading frame (ORF). Further examination of the DNA sequence in the region 5' to this indicated that there was a second potential ORF that could be brought in frame by the splicing out of a potential intron. A cDNA was cloned by hybridization of the region containing these two open reading frames, to a *S. pombe* cDNA library. Two clones were obtained, the longer of which was sequenced and was found to be 1073 base pairs long. The cDNA encoded a predicted protein of 331 amino acids long. A comparison between the genomic and cDNA sequences indicated that the gene identified on the 2.4kb *HindIII* fragment did indeed have a 44 base pair intron at the 5' end of the gene.

The putative protein encoded by the cDNA was compared to the protein sequences stored in the National Bethesda Research Foundation (PIR) and University of Geneva (SWISSPROT) protein sequence databases. Striking similarity with DNA methyltransferases specific for the 5-position of cytosine was noticed. The structural organization of the cnd1 protein sequence is very similar to these enzymes. Its predicted size of 38kd is within the range found for known bacterial m^5C-MTases. The ten conserved domains are present in the correct order and the spacing between them shows no atypical features. The cnd1 protein does not have the large N-terminal extension found in the mouse, human and *Arabidopsis* enzymes and is therefore more similar to the bacterial enzymes in its overall structure.

A schematic representation of the cnd1 sequence in comparison with a selection of enzymes is shown (Figure 1.1). The consensus sequence is based on an initial comparison of 13 m^5C-MTases (Posfai *et al.*, 1989). Since this initial analysis, the total number of these m^5C-MTases has risen to over 50 and most of them conform well to this consensus although only a small number of amino acids are absolutely invariant. Of the enzymes shown, M.MspI, M.HhaI and M.EcoRII conform well to the consensus (Lin *et al.*, 1989; Caserta *et al.*, 1987; Som *et al.*, 1987); these enzymes were part of the original analysis on which the consensus was based (Posfai *et al.*, 1989). The mouse sequence is shown as a representative of a eukaryotic m^5C-MTase and M.SssI and M.HgaI-2 are shown as examples of m^5C-MTases whose amino acid sequences deviate from the consensus at a number of points (Renbaum *et al.*, 1990; Sugisaki *et al.*, 1991). The cnd1 protein conforms reasonably well to the consensus, especially in motifs II, III, VI, VII, IX and X (Figure 1.1).

The most striking difference between cnd1 and all the other m^5C-MTases cloned to date occurs in region IV. In every other m^5C-MTase there is a proline-cysteine dipeptide in this motif that is known to be part of the catalytic site. The cnd1 protein, however, has a serine-cysteine dipeptide at this position. In addition to this change in region IV, there are a few differences in the cnd1 sequence at positions that correspond to residues that are usually invariant, most of these changes however, are conservative. For example, tyrosine-13 (region I), is found instead of the usually invariant phenylalanine, threonine-85 (region IV), is found instead of a serine, alanine-100 (region V), instead of leucine and asparagine-161 (region VIII) instead of glutamine.

The discovery of a m^5C-MTase homologue in fission yeast was surprising given the lack of detectable DNA methylation in this organism.

However, its association with a temperature-sensitive lethal mutation implied that there might be an essential function for DNA methylation in *S. pombe*. Based on these assumptions, further investigations into the possible role of the *cnd1*$^{+}$ gene were carried out and will be described in this thesis.

Figure 1.1 Amino acid sequence alignments to compare a selection of m^5C-MTases to the sequence of cnd1.
The enzymes included in this alignment are indicated at the left hand side (cnd1 is shown at the bottom of each block). The start point of each block within its particular protein sequence is indicated numerically. The borders of the conserved regions are as described (Posfai *et al.*, 1989). Blocks I, IV, VI, VIII, IX and X are considered to be highly conserved between m^5C-MTases whereas blocks II, III, V and VII are considered moderately conserved. The consensus sequence for m^5C-MTases is shown below each block. Bold type is used to indicate the agreement of a particular residue with the consensus. A further illustration of how cnd1 agrees with the consensus is also shown. * denotes identity and ^ denotes a conservative change between cnd1 and the consensus. Amino acids were grouped as follows for the purpose of describing a conservative change: (D,E), (S,T), (N,Q), (K,R), (F,Y), (A,V,L,I,M,C,W), (G,P).

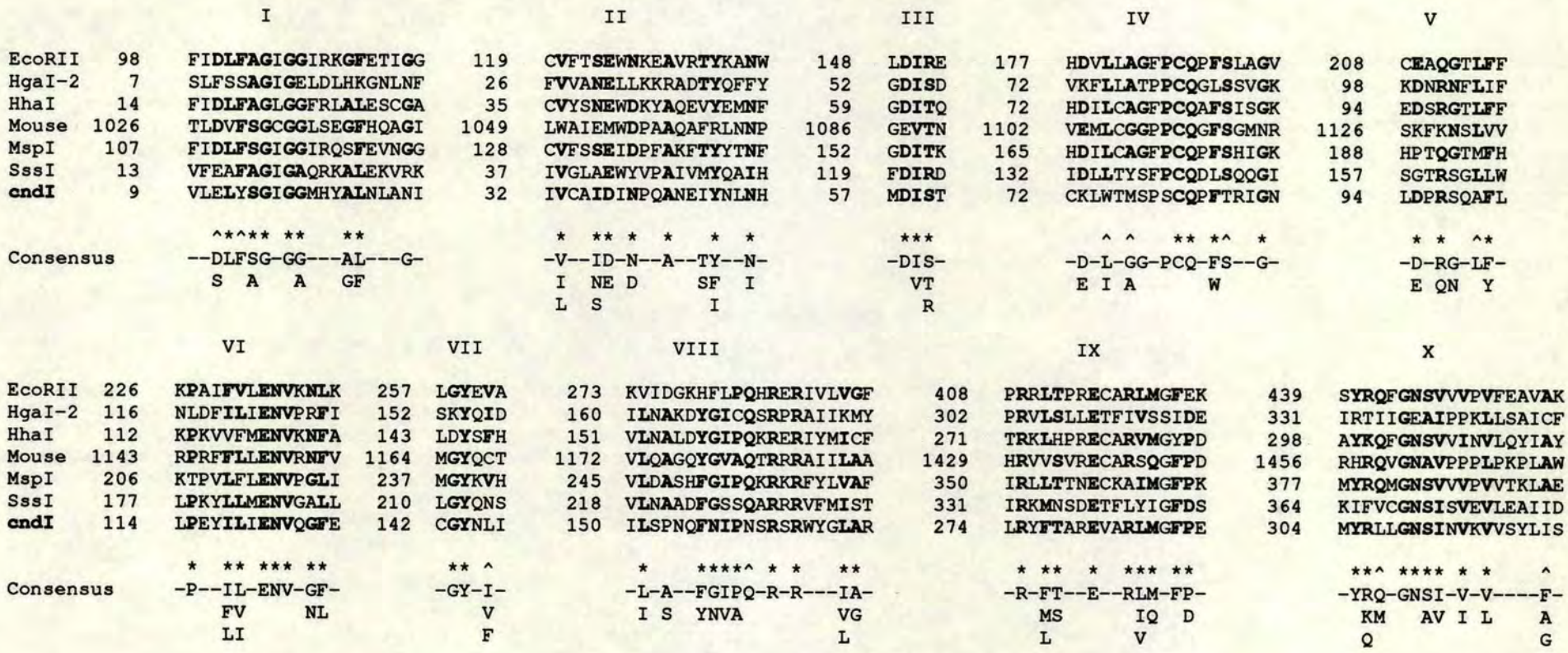


Figure 1.1 Amino acid alignments to compare a selection of m^5C-MTases to the sequence of cnd1.

Chapter 2 Materials and Methods

2.1 Commonly used reagents and buffers

Most of the methods used in this work were based on those described in Sambrook *et al.*, (1989). All chemicals were of analytical grade and were bought from Sigma, British Drug houses (BDH), Gibco-BRL, Fisons or Pharmacia. H_2O refers to distilled autoclaved water. Eppendorf tube refers to a 1.5ml centrifuge tube. High speed centrifugation in eppendorf tubes refers to an eppendorf centrifuge 5415C at 14000rpm.

Tris.HCl

Tris base (tris[hydroxymethyl]aminomethane, Sigma T-1503) was dissolved in H_2O and the pH adjusted to the required value by addition of HCl. H_2O was added to give a 1M stock solution.

EDTA

A stock solution of 0.5M EDTA (ethylenediaminetetramino acid, di sodium salt, Fisons D-0452) was made by dissolving solid EDTA in H_2O, adjusting the pH to 8.0 with NaOH, and adding water to the required volume.

TE

A buffer solution, suitable as a routine DNA solvent, consisting of 10mM Tris.HCl (pH 7.5) and 1mM EDTA.

Phenol

Phenol (Fisons, P-2360) was pre-equilibrated with 1M Tris.HCl (pH 7.5), followed by TE buffer. To retard oxidation of the phenol solution 0.1% (v/v) hydroxyquinoline (Sigma, H-5876) was added and it was stored at 4°C in the dark. For work with RNA phenol pre-equilibrated with H_2O (Rathburn 3024) was used.

Chloroform

Chloroform (BDH 10077) refers to a 24:1 (V/V) mixture of chloroform and isoamyl alcohol (Sigma, I-1381).

Sodium acetate

Sodium acetate was dissolved in H_2O, the pH adjusted to 5.2 with HOAc and H_2O added to a final concentration of 3M.

Ethidium Bromide

Ethidium bromide (Sigma, E-875) was dissolved as a stock solution of 10mg ml^{-1} in H_2O and stored at 4°C in the dark.

Loading buffer

10 x loading buffer for gel electrophoresis of nucleic acids was prepared and stored at room temperature:

40% glycerol
1% (w/v) orange G (Sigma, 0-1625)
10mM EDTA

TBE

TBE was routinely made up as a 10 X stock solution and stored at room temperature:

	stock	
0.89 M Tris base		108g
0.89 M boric acid		55g
20mM EDTA	0.5M	40ml
H_2O		to 1 litre

TAE

TAE was made up as a stock solution of 50 x and stored at room temperature:

	stock	
2M Tris base		242g
HOAc		57.1ml
50mM EDTA	0.5M	100ml
H_2O		to 1 litre

DEPC.H_2O

RNAse free water, suitable for dissolving RNA was made by adding 0.1% (v/v) diethyl pyrocarbonate (Sigma, D-5758) to H_2O, mixing for 5 minutes at room temperature and autoclaving.

2.2 Nucleic acid manipulations

i Dissolving and storage

All DNA and RNA was dissolved in TE or H_2O, depending on the subsequent use of the nucleic acid. Nucleic acids were routinely stored as solutions at -20ºC.

ii Extraction with phenol-chloroform

Proteins were removed from solutions containing DNA and RNA by extraction with an equal volume of a 1:1 mixture of phenol and chloroform. Traces of phenol were removed by a further extraction with chloroform. Extractions were carried out by adding a volume of phenol-chloroform equal to that of the nucleic acid solution. The two solutions were mixed thoroughly by vortexing to form an emulsion and separated by centrifugation, routinely 5 minutes, at high speed at room temperature. The aqueous phase was then transferred to a fresh tube carefully avoiding protein at the interface of the two phases and the nucleic acids recovered by precipitation.

iii Precipitation of nucleic acids

Both DNA and RNA were precipitated by the addition of a 1/10th volume of 3M sodium acetate (pH 5.2) followed by 2.5 volumes of absolute ethanol. The solution was mixed thoroughly by vortexing, frozen on dry ice for 5-10 minutes or -20ºC for 30 minutes and the nucleic acids pelleted by centrifugation at high speed at 4ºC. For the precipitation of small amounts of DNA, 1µl of carrier (glycogen 20mgml^{-1}, Boehringer 901393) was added before freezing.

To remove small amounts of salt remaining after the precipitation the

nucleic acid pellet was washed in 100µl of 70% (v/v) ethanol. The pellet was vacuum dried and dissolved in an appropriate volume of H_2O or TE buffer.

iv Quantification

DNA and RNA quality and amount was assayed by two methods: spectrophotometry and ethidium bromide staining after gel electrophoresis. The first method involves taking an absorbance reading at wavelengths 260 and 280nm. The reading at 260nm allows calculation of the concentration of nucleic acid in the sample. An A_{260} of 1.0 corresponds to approximately 50mgml^{-1} for double-stranded DNA, 40mgml^{-1} for RNA and 33mgml^{-1} for single-stranded oligonucleotides. The ratio between the readings at 260nm and 280nm provides an estimate of purity of the nucleic acid. Pure preparations of DNA and RNA have ratios of 1.8 and 2.0 respectively. Any ratios less than these values indicate protein and/or phenol contamination of the sample.

The second method utilizes ethidium bromide which specifically and proportionately binds to DNA and fluoresces under UV light. The nucleic acid sample is run on an appropriate gel containing ethidium bromide at a concentration of 0.5mgml^{-1} in parallel with a nucleic acid of known amount (usually standard molecular weight markers or λ DNA) and visualized on a trans-illuminator. A rough estimate of DNA amount can be gauged from the relative staining intensities of the bands.

v Plasmid vectors

pUC based plasmids

The pUC based plasmids are used for the manipulation of foreign genes in *E. coli* (Vieira and Messing 1982). They contain the pBR322 derived ampicillin resistance gene and origin of replication and coding information for the first 146 amino acids of the *lacZ* gene of *E. coli*. A polylinker is inserted into the *lacZ* coding region. This gene encodes the enzyme β-galactosidase. Non-recombinant plasmids, in the appropriate host cells, are able to synthesize this enzyme which breaks down X-gal to release a blue pigmented derivative. In recombinant plasmids the lacZ gene is interrupted by foreign DNA resulting in colourless colonies. The vectors pTZ18/19R (Pharmacia, 27-4984/6-01) are similar to pUC plasmids but contain in

addition, the f1 origin of replication. The plasmid pBluescript KS$^-$ (Stratagene, 212208) has a more extensive polylinker than pTZ plasmids and was used extensively throughout this project.

pIRT2

Plasmid pIRT2 (Hindley *et al.*, 1987; Figure 2.1) consists of the *LEU2* gene and the fission yeast autonomous replication sequence (*ars1*). It also contains the intergenic region of bacteriophage M13, enabling it to replicate to produce single-stranded DNA template suitable for DNA sequence analysis by the dideoxynucleotide method. It contains unique restriction enzyme sites for *SphI, PstI, SalI, BamHI, SmaI* and *SacI* in a polylinker.

pREP1/ pREP41/ pREP81

This family of vectors (Figure 2.2 and Table 2.1) is based on the thiamine repressible *nmt1* promoter (Maundrell, 1990). The promoter and polyadenylation signal of the *nmt1* gene are used to permit thiamine-mediated control of transcription of cloned genes (Maundrell, 1993). The pREP vectors contain the *LEU2* gene and the *S. pombe* ars origin of replication. They are maintained extrachromosomally at a high copy number. The polylinker contains *NdeI, SalI, BamHI* and *SmaI*. In minimal medium with no exogenous thiamine, the *nmt1* promoter is fully derepressed. Thiamine added to the medium is sequestered in the cells resulting in a rapid rise in the intracellular thiamine concentration and concomitant repression of *nmt1* transcription (Tommasino and Maundrell, 1991).

The pREP41 and pREP81 vectors are identical to pREP1 except that they carry TATA box mutations which result in lower expression from the *nmt1* promoter without affecting the repression by thiamine, or position of the transcription start point (Basi *et al.*, 1993).

TABLE 2.1 (adapted from Basi *et al.*, 1993)

Characteristics of the pREP1, pREP41 and pREP81 vectors and their relative promoter strengths under inducing and repressing conditions

Vector	TATA box	[a]Relative promoter activity	
		-thiamine	+[b]thiamine
pREP1	ATATATAAA	80	1
pREP41	ATAAA	12	0.06
pREP81	AT	1	0.004

[a]Relative promoter activity was measured as CAT activity. Data are normalized to the activity of the wild type promoter determined in cells cultured in the presence of thiamine (=1). For comparison, the relative values of the *adh* and *SV40* promoters are 12 and 2.5 respectively, when assayed either in the presence or absence of thiamine.
[b] Thiamine was added to the media to give a final concentration of 4μM. Thiamine is made up in H_2O (4mM stock solution) and is kept at room temperature in the dark.

pDB262

pDB262 (Figure 2.3) is capable of autonomous replication in both *E. coli* and *S.pombe* (Wright *et al.*, 1986). Cloning fragments into either the *HindIII* or *BclI* sites inactivates the λcI repressor gene product which allows expression of the tetracycline resistance gene, under the control of the bacteriophage λPr promoter. This enables positive selection of recombinant plasmids. Intact pDB262 does not express antibiotic resistance but can be selected in a *leu*$^-$ strain of *E. coli*. The presence of the *LEU2* gene permits selection for the plasmid in fission yeast.

pET6H

This plasmid is based on pET11D (Novagen; Figure 2.4).

Oligonucleotides with the sequence shown below were annealed and cloned into the *NcoI* and *BamHI* sites (C.H.Hu unpublished).

```
5' CATGCACCATCACCATCACCATGCCATGGAGCTCCCCGGGACTAGTG
       GTGGTAGTGGTAGTGGTACGGTACCTCGAGGGCCCTGATCACCTAG 5'
```

The resulting plasmid is called pET6H. In pET6H there are six histidines immediately downstream of the initiation methionine (indicated by an arrow), followed by novel sites for *NcoI*, *SstI*, *SmaI*, *SpeI* and the *BamHI* site of the original pET11D vector. The *NcoI* site at the initiating methionine is destroyed. The relevant sequences in pET6H are shown below.

start Met ↘

```
            M  H  H  H  H  H  H  A  M  E  L  P  G  L  V
GGAGATATACCATGCACCATCACCATCACCATGCCATGGAGCTCCCCGGACTAGTGGATCC
CCTCTATATGGTACGTGGTAGTGGTAGTGGTACGGTACCTCGAGGGCCCTGATCACCTAGG
              6 HISTIDINES    NcoI  SstI SmaI  SpeI  BamHI
```

pET6H is derived from pBR322 and can be used to express protein from a gene of interest under the control of the T7 promoter. The plasmid also contains the T7 transcription termination sequence. The stretch of consecutive histidine residues allows expressed protein to be purified using metal chelation chromatography.

Protein production is carried out in the strain *BL21 DE3 pLysS*. This strain contains a chromosomal copy of the T7 RNA polymerase gene. It is a lysogen of bacteriophage λDE3 which contains the polymerase gene under the control of the inducible lacUV5 promoter. Addition of IPTG to a growing culture induces the polymerase which in turn transcribes the target DNA in the plasmid. *BL21 (DE3)* also lacks the *ompT* outer membrane protease that can degrade some proteins during purification. This strain also contains the plasmid pLysS that provides a small amount of T7 lysozyme. The T7 lysozyme has two functions. Firstly, as a natural inhibitor of T7 RNA polymerase, it stabilizes target plasmids by decreasing the basal activity of T7 RNA polymerase, but it does not prevent induction of high levels of target proteins. Second it cleaves a bond in the peptidoglycan layer of the *E. coli* cell wall and allows cells to be lysed under mild conditions, such as freeze-thaw or the addition of 0.1% Triton X-100.

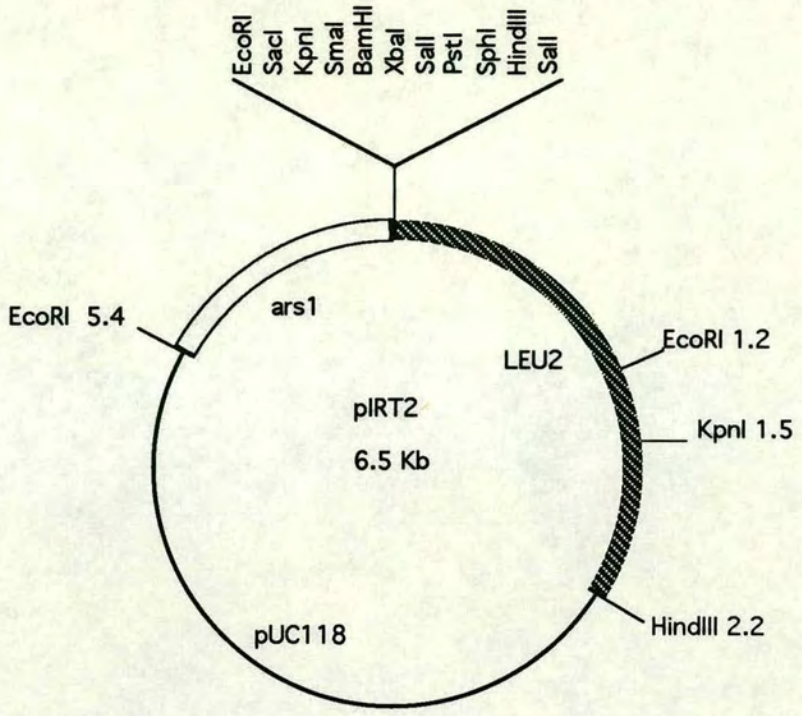


Figure 2.1 pIRT2

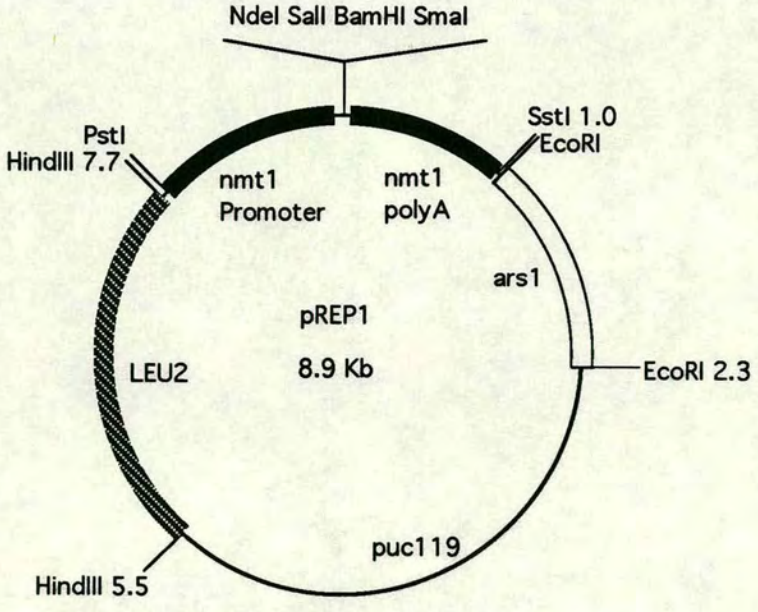


Figure 2.2 pREP1

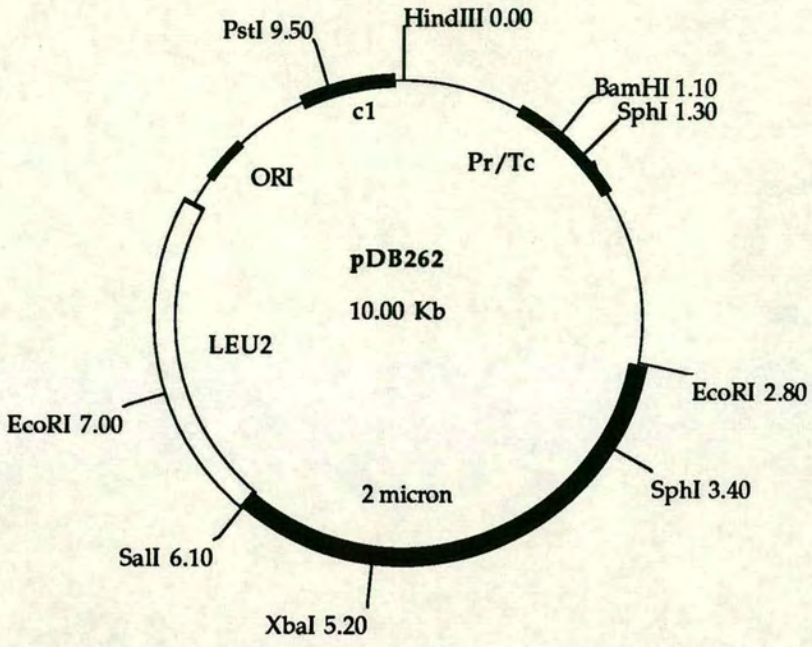


Figure 2.3 pDB262

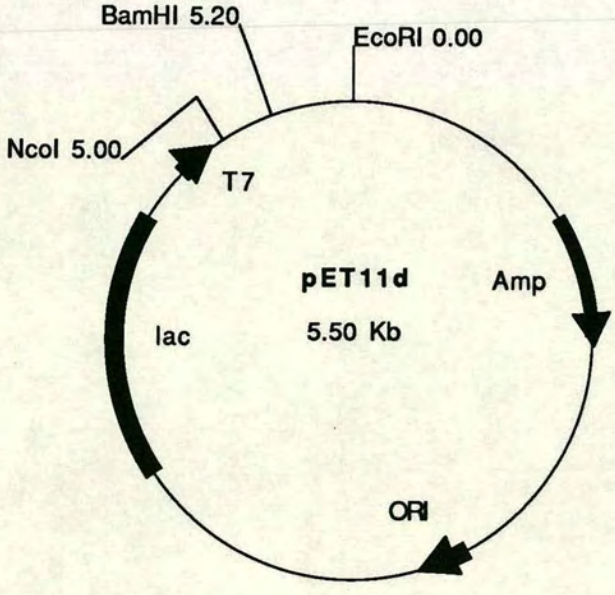


Figure 2.4 pET11d

2.3 Molecular analysis of nucleic acids

i Restriction enzyme analysis

Restriction enzymes were used, as recommended by the manufacturers, to cut DNA. The Boehringer restriction enzyme buffer system was used. Usually 1-10µg of DNA was incubated with 10 units of restriction enzyme and appropriate buffer in a final volume of 20-30µl at 37ºC for 1-15 hours.

ii Dephosphorylation of DNA

In order to prevent self ligation of vectors, the 5' phosphate was removed by alkaline phosphatase. Vector DNA was first cleaved with the appropriate restriction enzyme and then phenol-chloroform extracted followed by ethanol precipitation. The pellet was resuspended in TE and treated with bacterial alkaline phosphatase (Gibco, 18011-015) according to the manufacturer's instructions.

iii Ligations

Ligation of DNA with cohesive termini, was performed using T_4 DNA ligase (Boehringer, 862509) and ligation buffer. For blunt-ended termini, concentrated T_4 ligase (New England Biolabs, 202CS) and ligation buffer were used. Routinely between 1-10µg of DNA was ligated with 0.5µl of ligase in a total volume of 20µl. Nonidet-P40 was added to a final concentration of 0.1% to stimulate ligation efficiency. Ligation reactions were incubated for 16 hours at 16ºC.

Cohesive termini (with a 5' overhang) were converted to blunt ended termini by treatment with the Klenow fragment of DNA polymerase I. DNA was incubated at room temperature in the presence of 1 unit of Klenow fragment (Boehringer Mannheim, 100840), 50mM $MgCl_2$ and 0.2mM dNTPs for 30 minutes. The reaction was stopped by the addition of EDTA to 25mM. The DNA was then phenol-chloroform extracted and ethanol precipitated before further manipulations.

iv Agarose gel electrophoresis of DNA

Agarose gels were used to resolve DNA fragments from 100bp to 20kb using concentrations of agarose from 0.8 to 2.5% (w/v) in 1 x TAE buffer. The appropriate amount of agarose (Boehringer, 1444964) was dissolved in 10 x TAE buffer by boiling and upon cooling, ethidium bromide was added (to give a concentration of $0.3\mu g\mu l^{-1}$).

Samples were run for variable times at different voltages depending on the resolution required. A DNA marker was also used such as lambda DNA digested with *HindIII* or a 1kb ladder (Gibco-BRL, 5615SA).

v Purification of DNA fragments

DNA fragments between 500bp and 12kb were isolated from agarose gels by spinning through siliconized glass wool. The glass wool (Sigma, G-8389) was soaked in a 5% solution of trimethylchlorosilane (BDH, 3314640) and this was allowed to evaporate. The wool was then washed thoroughly in H_20 and baked dry at 80°C for 2 hours.

A hole was made in the bottom of a 0.5ml eppendorf tube and a plug was made with the glass wool. A hole was also made in the lid of the tube. The gel slice containing the fragment of interest was inserted into the tube and the tube placed inside a 1.5ml eppendorf tube. This was spun at 6000rpm for 10 minutes. The DNA solution collected in the 1.5ml tube was then phenol-chloroform extracted and ethanol precipitated. If the DNA was to be used for random prime labelling, it was used directly without extraction and precipitation.

vi Nearest neighbour analysis and thin layer chromatography (TLC)

1. Fission yeast DNA (2µg) was digested with *BstNI* for 7 hours at 60°C.
2. Blunt ends were created at the resulting overhanging ends using the Klenow fragment of DNA polymerase I in the presence of [α-^{32}P] dATP (section 2.4 i).
3. The DNA was ethanol precipitated to get rid of unincorporated counts and was resuspended in the following:

	stock	
20mM $CaCl_2$	0.1M	4µl
[1]micrococcal nuclease		1.5µl
[2]phosphodiesterase (PDE)		4µl
100mM Tris.HCl pH 8.5	1M	2µl
H_2O		8.5µl

[1]Micrococcal nuclease (Sigma, N 3755)
[2]PDE from calf spleen (Boehringer, 108260)

4. The mixture was incubated for 4 hours at 37°C.
5. TLC plates (Sigma, T-5770, 250µM layer) were washed 3 times in methanol and allowed to dry.
6. The DNA was spotted directly onto the plate along with 10µg of each nucleotide standard (ammonium salts dissolved in H_2O, $10\mu g\mu l^{-1}$ Sigma, D6897, D1417, D6772, D7645).
7. The first dimension was run for 12-15 hours in solvent 1. The solvent had previously been allowed to equilibrate in the tank to aid saturation. The plate was placed in the tank such that the solvent reached 0.5cm below where the DNA mix had been loaded.
8. The plates were allowed to dry and then washed 3 times in methanol and allowed to dry again.
9. The 2nd dimension was run for 4 hours, after which the plates were dried.
10. Labelled nucleotides were detected by autoradiography and compared to the position of the standards detected by autoradiography and compared to the positions of the standards detected by fluorescence.

1st dimension solvent

Isobutyric acid (Sigma, I 1754)	189ml
H_2O	51ml
NH_3 (28.5% NH_3 in H_2O; Sigma, A6899)	10ml

2nd dimension solvent

saturated $(NH_4)SO_4$	200ml
Isopropanol	5ml
sodium acetate pH 5.2 (1M)	45ml

vii Gel electrophoresis of RNA

RNA was separated using a formaldehyde gel (Golberg, 1980). Gloves were worn during all procedures with RNA to prevent RNAse contamination.

Formaldehyde gels

1. 1.0g agarose was dissolved in 63ml H_2O and 16ml 5 x [1]MNE buffer by boiling: this was cooled to 60°C, 17ml of 38% (w/v) solution of formaldehyde (BDH, 28121) added, and this was poured immediately into a horizontal gel mold.
2. Between 5-20µg of fission yeast total RNA was run per lane: to each sample, 15µl of [2]RNA buffer and 1µl of 0.5mg ml^{-1} ethidium bromide was added. The RNA was incubated at 60°C for 5 minutes and loaded onto the gel.
3. Gels were run in 1xMNE buffer with an applied voltage of 10V cm^{-1} for 3 hours.

[1]5 x MNE buffer

0.12M MOPS (Sigma, M-1254)	13.08g
40mM NaOAc	1.03g
5mM EDTA	0.19g
DEPC.H_2O	to 500ml

Adjusted to pH 7 with NaOH and stored at 4°C in the dark.

[2]RNA buffer

formaldehyde	600µl
formamide (BDH, 28241)	200µl
MNE Buffer	240µl
DEPC.H_2O	160µl

2.4 Radio-labelling of DNA fragments

Radio-labelled fragments of DNA were used as probes for Northern and Southern blotting (section 2.5). All of the probes that were used in this project were labelled using the random prime method apart from the LTR probe of the *S. pombe* transposable element which was end-labelled due to its small size (348bp).

i Random prime labelling

This method is based on that described by Feinberg and Vogelstein (1983). The method uses random hexanucleotides which bind to the DNA fragment and initiate DNA polymerase reactions with the Klenow fragment of DNA polymerase I. One of the nucleotides incorporated is radio-labelled.

1. Digested DNA was run through a 0.8% agarose gel, and visualized by ethidium bromide staining.
2. The fragment of interest was then excised and spun through siliconized glass wool (section 2.3 v).
3. The resulting DNA was tested by running on an agarose gel alongside a known quantity of DNA to assess the concentration of the purified fragment. Approximately 50ng of DNA was used in a labelling reaction.
4. The labelling reaction consisted of:

[1]OLB	10µl
DNA (20-50ng)	2-17µl
[α^{32}P] dCTP (Amersham 10µCiµl^{-1})	3µl
Klenow (Cambio, 022366500)	1µl
H_2O	to 50µl

This reaction was incubated for 30 minutes at 37°C.

5. The reaction was stopped and the DNA precipitated by the addition of the following:

0.5M EDTA	2µl
2M NaOAc	5µl
glycogen 20mgml^{-1}	1µl

100% ethanol 125µl

The DNA was washed after precipitation in 70% ethanol and vacuum dried.

6. Labelled DNA was resuspended in 100µl TE and boiled for 5 minutes before adding to the hybridization solution.

[1]**OLB** OLB was made from the following components :

O 18µl of β-mercaptoethanol and 5µl each of 100mM dATP, dTTP and dGTP (Boehringer, 1277049) were added to 1ml of 1.25M Tris.HCl (pH 8.0) and 0.125M $MgCl_2$.

L 2M Hepes (Sigma H-3375) titrated to pH 6.6 with 4M NaOH

B Hexa-deoxyribonucleotides at 90 OD units ml^{-1} (Boehringer 1277081).

Solutions O:L:B: were mixed at a ratio of 100: 250: 150 to make OLB. This was stored as 50µl aliquots at -20ºC.

ii End-labelling

This method is based on that of Sambrook *et al.*, (1989). End labelling uses bacteriophage T_4 polynucleotide kinase (PNK) which catalyzes the transfer of the γ-phosphate of ATP to a 5' terminus of DNA or RNA. [γ-^{32}P] ATP is added in equimolar amounts to the number of ends of nucleic acid in the presence of PNK. This results in terminal γ-phosphates being replaced by [γ-^{32}P] phosphates and so the DNA fragment is radio-labelled.

1. The following reagents were mixed together:

DNA	10pmoles	
[1]PNK buffer	1µl	
[γ-^{32}P] ATP	3µl	(Amersham, 10µCiµl^{-1})
T_4PNK	1µl	(Gibco BRL, 10Uµl^{-1})
H_2O	to 10µl	

2. The reaction mixture was incubated for 30 minutes at 37ºC and stopped by incubating at 70ºC for 5 minutes. Labelled DNA was

separated from unincorporated counts by precipitation.

[1] PNK buffer

	stock	1ml
0.5M Tris.HCl pH 7.5	1M	500µl
100mM $MgCl_2$	1M	100µl
50mM DTT	1M	50µl
1mM spermidine	1M	1µl
1mM EDTA	0.5M	2µl
H_2O		347µl

2.5 Filter hybridization of nucleic acids

Southern blot analysis

i Transfer

DNA was transferred from agarose gels to a nylon membrane (Hybond-N^+, Amersham RPN 203B) using the method of Southern (1975).

1. DNA digested with the appropriate restriction enzyme was run on an agarose gel. The gel was then gently shaken in 0.5mgml^{-1} ethidium bromide (diluted in 1 x TAE) for 30 minutes and then destained by shaking in 1 x TAE for 30 minutes.
2. The gel was photographed, exposed to UV light for 5 minutes, and rinsed with H_2O.
3. The gel was denatured in 0.4M NaOH and 1.5M NaCl for 30 minutes, and then neutralized in 0.5M Tris.HCl pH 7.5 and 3M NaCl. The transfer was performed in 2 x SSC and the membrane and 3MM paper were also soaked in 2 x SSC before setting up the blot.
4. The gel was laid onto a piece of clingfilm and surrounded by strips of parafilm. The gel was then overlaid with a piece of membrane soaked in 0.4M NaOH and 3 pieces of 3MM filter paper (also soaked in 0.4M NaOH) and a stack of paper towels. Transfer was allowed to proceed overnight.
5. After transfer the filter was labelled and rinsed in 2 x SSC and the DNA crosslinked to the membrane by exposure to UV.

ii Hybridization

Filters were prehybridized in 20ml of [1]hybridization solution (a modification of the buffer used by Church and Gilbert, 1984) in a Techne hybridization oven for 1 hour. Hybridizations were usually carried out at 65ºC overnight. The filters were washed 3 times in 0.2 x SSC and 0.1% SDS at 65ºC. The filters were sealed in saran wrap before autoradiography.

[1] **Hybridization solution**

	stock	
0.5M Na_2HPO_4 (pH 7.2)	1M	10ml
7% SDS	20%	7ml
1mM EDTA	0.5M	40ml
0.5% dry milk	0.5g in 10ml	2ml
H_2O		to 20ml

iii Autoradiography

Autoradiography was carried out using Kodak X-AR S autoradiograph film in cassettes which had tungsten intensifying screens. Filters were exposed for various times at -70ºC. ^{35}S and very radioactive filters were exposed at room temperature. Films were developed using an X-OGRAPH compact X2 automatic processing machine.

iv Stripping

Filters were stripped using boiling water to which SDS was then added to a final concentration of 0.5%. The filter was shaken gently at room temperature until the solution had cooled.

Northern Blot analysis

v Transfer

Northern blotting was performed using the capillary method onto genescreen membrane (NEN research products), following the manufacturer's instructions.

1. Genescreen membrane was cut to slightly larger than gel size, and soaked in 0.025M Na_2HPO_4/NaH_2PO_4 (pH 6.5)[1] for 20 minutes; the RNA gel was soaked in this buffer for the same time.
2. Ten pieces of filter paper (Whatman 3MM) were wetted with 0.025M Na_2HPO_4/NaH_2PO_4 (pH 6.5) and placed on an elevated glass plate so that the ends formed a wick.
3. The gel was placed on filter paper, avoiding air bubbles and gel spacers placed alongside each side of the gel.
4. Genescreen membrane was placed on the gel ensuring there were no air bubbles.
5. Five pieces of Whatman paper cut to the same size as the gel were placed on top.
6. Twenty absorbent hand towels were placed on the filter paper and a weight placed on top.
7. RNA was transferred for at least 20 hours.
8. Towels and Whatman paper were discarded and the membrane was washed in 0.025M Na_2HPO_4/NaH_2PO_4 (pH 6.5) buffer for 20 minutes with gentle agitation.
9. The membrane was dried on filter paper at room temperature and then baked at 100ºC for 3 hours; membranes could be stored at this stage.

[1] 10mM Na_2HPO_4/NaH_2PO_4 (pH 6.5) was made by adding 3.9ml of 1M NaH_2PO_4 and 6.1ml of 1M Na_2HPO_4 in 90ml of H_2O.

vi Hybridization

1. Membranes were prehybridized in [1]hybridization buffer for 4 hours at 42ºC in a Techne hybridization oven.
2. Radio-labelled probe was boiled for 5 minutes and chilled on ice for 10 minutes: this was added to the hybridization buffer and hybridized for 16 hours at 42ºC.
3. Hybridization buffer was discarded and the membrane washed in the oven as follows:
   twice with 100ml of 2 x SSC[2] at 42ºC for 5 minutes;
   twice with 100ml of 2 x SSC plus 1.0% SDS at 65ºC for 30 minutes;

twice with 100ml of 0.1 x SSC at room temperature for 30 minutes.

4. The membrane was never allowed to become dry at this stage as the probe became covalently bound to the RNA preventing further hybridizations.
5. The membrane was processed for autoradiography.

[1]Hybridization buffer:
10ml de-ionized formamide, 4ml P buffer and 4ml 50% (w/v) dextran sulphate were added together, mixed and incubated at 42°C for 10 minutes. 1.16g NaCl was added, mixed and incubated at 42°C for a further 10 minutes until it had dissolved. 200µl of denatured salmon sperm DNA was added and mixed.

P buffer:
1.0% (w/v) bovine serum albumin (globulin free fraction, Sigma, A7638), 1.0% (w/v) polyvinyl-pyrrolidone (Sigma, P-5288), 1.0% (w/v) ficoll (Sigma, F-2637), 250mM Tris.HCl (pH 7.5), 0.5% (w/v) tetra-sodium pyrophosphate (BDH, 10261) and 5% (w/v) SDS.

50% (w/v) dextran sulphate (sodium salt, Pharmacia, 17-0340-01) solution. De-ionized formamide was prepared by mixing 100ml of formamide (BDH, 28241) with 5g mixed bead resin (AG 501-X8 [D]) at room temperature for 30 minutes. The solution was then filtered to remove the resin. This was stored in aliquots at -20°C.
Salmon sperm DNA (Sigma, D-1626): this was dissolved at a concentration of $10mgml^{-1}$ in H_2O and sonicated until the solution was liquid. Prior to use, DNA was denatured by boiling for 10 minutes and then chilled on ice for a further 10 minutes.
[2] SSC was made as a 20 x stock as follows:

3M NaCl	876.6g
0.3M sodium citrate	441.15g
H_2O	to 5 litres

vii Autoradiography

Genescreen membranes were sealed in saranwrap and processed for autoradiography.

viii Stripping

Blots were stripped of radio-labelled probe, permitting further hybridizations with other probes.

1. Membranes were washed in 250ml of [1]strip solution for 2 hours at 70ºC in Techne hybridization ovens.
2. Membranes were air dried on Whatman paper (3MM) and stored.

[1]Strip solution

	stock	
5mM Tris.HCl pH 7.5	1M	5ml
0.2mM EDTA	0.5M	400µl
0.05% (w/v) Na pyrophosphate		500mg
0.002% (w/v) polyvinyl-pyrrolidone		20mg
0.002% (w/v) bovine serum albumin		20mg
0.002% (w/v) ficoll		20mg
H_2O		to 1 litre

2.6 Polymerase chain reaction

The polymerase chain reaction (PCR) was used for the amplification of DNA fragments for a wide number of uses. Amplification was performed over 30 cycles using 0.1µg of chromosomal DNA or 10-50ng of plasmid DNA with the Promega Taq polymerase and buffer system (M1861) in a Hybaid omnigene machine. Programmes were as follows:

94ºC for 1 minute (denaturation)
*55ºC for 1 minute (annealing)
72ºC for 1 minute (elongation)

After 30 cycles of this the reaction was heated to 72ºC for 10 minutes.
* different annealing temperatures were used according to the Tm values of the primers. Details of all the primers used throughout this project and the conditions for each individual reaction are shown in Appendices A and B. Each reaction mixture was made up as follows:

Template DNA	10-100ng
[1]primer 1 (25pmolμl^{-1})	1µl of 300ngμl^{-1}
primer 2 (25pmolμl^{-1})	1µl of 300ngμl^{-1}
dNTPs (10mM)	2µl
[2]10 x reaction buffer	10µl
Taq pol	0.5µl
H_2O	to 100µl

To prevent evaporation, each reaction mix was overlaid with 100µl of mineral oil (PCR grade Sigma, M-3516).

[1] The oligonucleotides used as primers in PCR reactions are listed in Appendices A and B. They were synthesized by either the ICRF oligonucleotide synthesis service or by Oswel DNA custom oligonucleotide synthesis, Dept. of chemistry, University of Edinburgh.
[2] For PCR of fragments of > 2kb in length, Thesit reaction buffer was used.

10 x Thesit buffer
300mM Tricine pH 8.4
20mM $MgCl_2$
50mM β-mercaptoethanol
0.1% gelatine (2% stock)
1% Triton X-100

Reverse transcriptase PCR

1. 1µg of fission yeast total RNA was made up to a volume of 15µl with DEPC.H_2O and 4µl of oligo dT (1µgμl^{-1}, ICRF) was added.
2. The RNA was denatured for 5 minutes at 70°C.
3. The RNA and oligo dT were allowed to anneal on ice for 3 minutes.
4. The following was then added to the mixture:
   0.5µl RNAsin (5 units, Promega, N2111)
   4µl 10mM [1]dNTPs
   2µl Reverse Transcriptase (RT, 200uμl^{-1}, Gibco, 18053-017)
   8µl RT buffer
   4µl 0.1M DTT

In addition to this, a duplicate sample without RT was set up.

5. 5µl was removed from this mixture and added to 0.5µl ^{32}P dCTP.
6. The mixture was incubated for 1h at 37°C.
7. The mixture was then heat inactivated for 10 minutes at 70°C.
8. 60µl of DEPC.H_2O was added to the main sample.
9. The radioactively labelled sample was ethanol precipitated, washed with 70% ethanol, dried and resuspended in TE. This was then run on an agarose gel. After running the sample the gel was placed on two pieces of DE81 paper with two pieces of 3MM Whatman paper below and dried at 80°C for 45 minutes. The gel was then exposed to Kodak film and processed as described (section 2.5 iii) to assay the extent of cDNA synthesis.
10. The PCR reaction was carried out essentially as described (section 2.6), using 1/10th of the RT reaction. Samples of plus and minus RT were used.

2.7 DNA sequencing

DNA was sequenced using the chain termination method of Sanger *et al.*, (1977). Two methods were used both using double stranded plasmid DNA but using different labels.

i. ^{35}S

ii. Fluorescent dyes

i Deletions

A series of unidirectional nested deletions of the suppressor of *cnd1-1* subcloned in pBluescript was made (section 2.2 v) using the method of Henikoff (1984). This method uses exonuclease III (exoIII), a 3' exonuclease which is only active on double stranded DNA. Blunt and 5' overhanging ends are susceptible to digestion while 3' overhanging ends of three or four bases in length are resistant to the enzyme.

It is therefore possible to prepare linearized DNA which has only one susceptible end next to the DNA fragment of interest by cutting the DNA with two restriction enzymes, one of which produces a susceptible end and the other a resistant end. Incubation of this DNA with exoIII under

controlled conditions of salt and temperature permits the progressive removal of nucleotides from one strand. By stopping the reaction at incremental times, a series of DNA pieces of varying length can be produced. The remaining single stranded DNA is removed using S1 nuclease, and the molecules are recircularized using T_4 DNA ligase. These nested deletions can then be sequenced using a common oligonucleotide primer, which hybridizes to a DNA sequence just internal to the resistant end.

Restriction enzyme digestion

To create a linearized molecule with resistant and sensitive ends to exoIII, 10µg of pB1-2 was digested with *KpnI* (to give the resistant end) and *SalI* (to give the sensitive end) and the digestion monitored by gel electrophoresis. The enzymes were heat inactivated.

Exo III digestion

ExoIII digestion was performed at 37ºC in 50mM NaCl to give deletions of 300 nucleotide incremental size. The DNA was already in Boehringer restriction enzyme buffer H (100mM NaCl). The reaction volume was adjusted so that the final concentration of NaCl was 50mM, exoIII and [1]exoIII buffer were added.

The mix was prewarmed to 37ºC and a 2µl "time zero" sample removed and added to 3µl of ice cold 2 x S1 buffer [2]. 1µl of exoIII (Gibco, 5100-8013SA), was added and further 2µl samples removed at 1 minute intervals and added to 3µl of stop buffer.

cut DNA in buffer H	20µl
ExoIII buffer	4µl
ExoIII	1µl
H_2O	15µl

After all timed samples had been removed, they were incubated simultaneously at room temperature for 30 minutes after which 1µl of stop solution (7mM EDTA) was added to each tube, mixed and incubated at 65ºC for 10 minutes. The deletions were analyzed by gel electrophoresis.

[1]10 x exonuclease buffer
0.66M Tris.HCl pH 8.0
66mM $MgCl_2$

[2]10 x S1 nuclease buffer
2M NaCl
0.5M NaOAc (pH 4.5)
10mM $ZnSO_4$
5% glycerol

S1 nuclease was diluted into this buffer so that the final reaction mixture contained 0.03 units of enzyme.

Recircularization and transformation

1. The remaining 3µl of each reaction mixture was ligated in a volume of 100µl to promote intramolecular ligation.
2. Each ligated sample was transformed into the bacterial strain *XL-1 Blue* (section 2.9 i).
3. Miniprep plasmid DNA was prepared (section 2.9 iii) and the DNA was sequenced. pB1-2 was sequenced using the ^{35}S sequencing method. Later sequencing in the project was done using an ABI automated sequencing machine.

ii ^{35}S Sequencing

This was carried out using the sequenase version 2.0 kit (USB).

1. 1-3µg of double stranded plasmid DNA was mixed with 10ng of primer in a volume of 20µl.
2. 20µl of 0.4M NaOH was added and the mixture was incubated for 5 minutes at room temperature.
3. 4µl of 2M NH_4OAc pH 4.5 and 80µl of ethanol were added.
4. This was precipitated, washed with 70% ethanol and dried.
5. The pellet was resuspended in:

H_2O	7µl
5 x sequenase buffer	2µl

The following was then added:

0.1M DTT	1µl
[1][α-^{35}S] dATP	0.5µl
labelling mix	2µl (USB kit stock diluted 1 in 8).
Diluted T7 polymerase	2µl (diluted to 1.5uµl^{-1}

[1][α-^{35}S] dATP: Amersham 37.0 MBq 1.00mCi; 10µCiµl^{-1}

6. This mixture was incubated for 5 minutes at room temperature.
7. 3.5µl was added to each 2.5µl termination mix (prewarmed at 37°C)
8. This mixture was incubated at 37°C for 5 minutes.
9. 4µl of formaldehyde stop mix was added and the samples left on ice until ready to run the gel.

Gel electrophoresis

Electrophoresis was carried out using the Bio-Rad sequencing system.

1. Plates were cleaned with a mild detergent, rinsed with distilled water and wiped with 70% ethanol.
2. Repelcote (a siliconizing treatment, BDH, 63216 4J) was applied to the dried plates and spacers inserted.
3. To 30µl of [1]sequencing mix, 150µl of freshly prepared 25% ammonium persulphate (Sigma, A-9164) and 150µl of TEMED (N,N,N',N'- tetramethylethylenediamine Bio-Rad, 61-0800) were added. This was poured over a wick in a casting tray, the gel plates were slotted in and the mixture allowed to set.
4. 100µl of ammonium persulphate and 100µl of TEMED were added to the rest of the gel mix, this was poured into the plates and allowed to set with a "sharks tooth" comb to form a loading well.
5. Gels were prerun for 30 minutes at 2000V in 1 x TBE buffer.
6. The wells were flushed out and the denatured samples (80°C for 3 minutes) loaded 4µl of sample was loaded per lane.
7. Gels were run at between 1600 and 2000V for 2.5 to 6 hours.

8. Gels were transferred to a sheet of 3MM Whatman filter paper, covered with cling film and dried on a vacuum gel dryer.
9. The gels were processed for autoradiography.

[1] Sequencing mix	stock	
7.6M urea		46g
6% Acrylamide (19:1)	30%	20ml
1 X TBE	10 X	10ml
H_2O		to 100ml

iii Automated sequencing

Double stranded plasmids were also sequenced using an Applied Biosystems automated sequencing machine 373A. This was done entirely according to the manufacturer's instructions. Cycle sequencing (denaturation, annealing and extension), was carried out in a PCR machine. Since only a single primer is used, DNA templates are linearly amplified. An ABI Tae Dyedeoxyterminator cycle sequencing kit was used (ABI 401113). Each of the dideoxynucleotides were labelled with a different fluorescent dye so that the four base reactions were performed simultaneously in one tube and were electrophoresed together in one gel lane. After the cycling stage the unincorporated dye terminators are removed by organic extraction.

2.8 Protein manipulations

i SDS-PAGE analysis of proteins

Discontinuous SDS-polyacrylamide gels were used for the separation of proteins under denaturing conditions. Gels were poured using the Mini-Protean II dual slab cell for miniature polyacrylamide gels (BioRad, 165-2940). Discontinuous polyacrylamide consists of a resolving (lower) gel and a stacking (upper) gel. The stacking gel acts to concentrate large sample volumes, resulting in a better gel resolution. Molecules are then completely separated in the resolving gel. Reagents and gel preparation described

below are based on the method described by Laemmli (1970) and on the instruction manual of the Bio-Rad Mini-Protean II system.

1. [1]Resolving gel monomer solution was prepared and poured. The monomer solution was immediately overlaid with water.
2. The resolving gel was allowed to polymerize for 45 minutes and then the water was poured off.
3. The [2]stacking gel monomer solution was prepared, poured immediately and the combs fitted. The stacking gel was allowed to polymerize for 45 minutes and the combs were removed.
4. The gel apparatus was fitted into the buffer tank and the tank filled with [3]Tris-glycine buffer.
5. Protein samples prepared in [4]SDS-loading buffer were loaded onto the gel. Cell pellets were resuspended in [5]8M urea loading buffer before loading. The samples were denatured at 95ºC for 3 minutes (in a heating block). As a molecular marker, pre-stained SDS-PAGE low range molecular weight markers (Bio-Rad, 161-0305) were used.
6. Gels were run at 200 volts for approximately 45 minutes.
7. After electrophoresis, the gels were stained with [6]Coomassie Brilliant Blue dye (Bio-Rad, 161-0400) for 5-15 minutes. The dye was removed and the gels rinsed briefly with [7]destain solution and left to destain overnight with a destain bag (MoBiTec).
8. Destained gels were placed on 2 pieces of Whatman paper, covered with saranwrap and dried at 80ºC under a vacuum for 1 hour.

[1]Resolving gel (10%)

30% Acrylamide (29:1) Bisacrylamide solution (NBL, 077234)	6.6ml
20% (w/v) SDS	100µl
1.5M Tris.HCl pH 8.8	5ml
[a]10% (w/v) Ammonium persulphate	125µl
[a]TEMED (N,N,N',N'-Tetramethylethylenediamine)	15µl
H_2O	8.1ml

[2]Stacking gel

30% Acrylamide (29:1) Bisacrylamide	1.6ml
0.5M Tris.HCl pH 6.8	2.5ml
20% (w/v) SDS	25µl
[a]10% ammonium persulphate	100µl
[a]TEMED	7µl
H_2O	5.8ml

[a] Ammonium persulphate and TEMED were added to the other components just before the mix was poured into the gel plates.

[3]Tris Glycine buffer

25mM Tris
250mM glycine pH 8.3
0.1% SDS

[4]SDS-loading buffer (2 X)	stock	
62.5mM Tris.HCl pH 6.8	0.5M	6.25ml
2% (w/v) SDS	20%	5ml
β-mercaptoethanol		1ml
10% glycerol	50%	10ml
0.1% bromophenol blue	1%	5ml
H_2O		to 50ml

[5]8M urea loading buffer	stock	
10mM NaPi pH 7.2	1M	200µl
1% (w/v) SDS	20%	1ml
8M urea		9.6g
0.1% Triton X-100	1%	2ml
1% β-mercaptoethanol		200µl
H_2O		to 20ml

[6]Coomassie Brilliant Blue dye solution

0.1% (w/v) Coomassie blue dye
50% methanol
10% acetic acid

[7]Destain solution

5% methanol

7% acetic acid

ii Purification of the cnd1 protein

Induction

1. The cnd cDNA was subcloned into the pET6H vector to give pET6Hcnd1 (details in Appendix B and section 2.2 v).
2. The pET6Hcnd1 plasmid was transformed into the strain *(DE3) pLysS* and transformants selected on plates containing 50μgml^{-1} ampicillin and 34μgml^{-1} chloramphenicol.
3. A 50ml preculture (L-broth containing ampicillin and chloramphenicol) was inoculated with a single transforming colony and allowed to grow overnight at 24ºC in an orbital shaker.
4. The following morning the preculture was diluted into 1 litre of L-broth with selection as before and allowed to grow to an OD_{600} of between 0.4 and 0.8 at 24ºC.
5. Isopropylthio-β-D-galactoside (IPTG) was added to the culture to give a final concentration of 0.4mM and the cells left in the shaker for a further 4 hours.

Preparation of cell extracts

6. Cells were pelleted in a Beckman JB-6 centrifuge for 20 minutes at 3.6K and 4ºC.
7. The cell pellet was resuspended in 25ml of [1]buffer A using a plastic pasteur pipette.
8. The cells were broken open by sonication using a Lucas Soniphore sonicator. The cells were placed in a 50ml beaker on ice and sonicated for 4 minutes, power 5 on the 40% duty cycle.
9. The sonicated sample was then left stirring at 4ºC for 1-2 hours to allow for solubilization of proteins.
10. Insoluble cellular material was pelleted at 16K for 30 minutes in the JA-20 rotor of a Beckman J2-20 centrifuge.

11. The pellet was resuspended in 8M urea loading buffer (section 2.8 i) to allow analysis of insoluble proteins by SDS-PAGE.
12. Soluble material (also analyzed by SDS-PAGE by the addition of 2 X SDS loading buffer [section 2.8 i]) was filtered through a 0.45μM membrane. The filtered sample was applied to the nickel agarose column.

Nickel-ion affinity chromatography

13. A slurry of nickel agarose (Novagen, 69670-2) was poured into a plastic column and allowed to pack under gravity to give a total of 1ml in volume. A 19 gauge needle was attached to the bottom to increase the flow rate.
14. The column was washed with 3ml of H_2O.
15. The column was charged by adding 5ml of 50mM $NiSO_4$ solution.
16. The column was equilibrated with 5ml of buffer A.
17. 25ml of cell extracts were added and allowed to flow through the column. The flow through was collected as one fraction.
18. The column was washed with 6ml of buffer A containing 10 or 20mM imidazole, collected as one fraction.
19. The column was washed with 6ml of buffer A containing 40mM imidazole collected as three 2ml fractions.
20. The majority of the cnd1 protein was eluted with 80 and 100mM imidazole (in buffer A). 6ml of buffer A containing 80mM imidazole was passed over the column and 3 fractions each of 2ml were collected. The same procedure was applied with 6ml of 100mM imidazole.
21. After use the column was washed with 6ml of buffer A and then stripped with 100mM EDTA before reuse.

[1]Buffer A-50mM NaCl

	stock	buffer A
10mM imidazole	1M	1ml
50mM NaCl	5M	1ml
10mM Tris.HCl pH 7.9	1M	1ml
10% glycerol	50%	20ml
0.1% Triton	20%	0.5ml
[a]0.5mM PMSF	100mM	0.5ml
[b]protease inhibitors	$5mgml^{-1}$	100µl of each
H_2O		to 100ml

Buffer for preparing cnd1 antigen for generation of polyclonal antisera is the same as buffer A except that 500mM NaCl was used.

Buffer B-6M urea to solubilize protein grown at 37°C

	stock	Buffer B
10mM Imidazole	1M	1ml
6M urea		36g
0.5M NaCl	5M	10ml
10mM Tris.HCl pH 7.9	1M	1ml
10% glycerol	50%	20ml
0.1% Triton	20%	500µl
[a]0.5mM PMSF	100mM	500µl
[b]protease inhibitors	$5mgml^{-1}$	100µl of each
H_2O		to 100ml

Buffer C-for the resuspension of dialyzed cnd1 pellet for antibody production

	stock	Buffer C
10% glycerol	50%	400µl
20mM Tris.HCl pH 7.9	1M	40µl
100mM NaCl	5M	20µl
H_2O		2ml

[a]added just before use
[b]pepstatin A (Sigma, P-4265), leupeptin (Sigma, L-2884), chymostatin (Sigma, C-7268) and antipain (Sigma, A-6271).

iii Methyltransferase assays

Methyltransferase activity was analyzed by detecting the transfer of ^{3}H-methyl groups from ^{3}H-AdoMet to a variety of substrates. The assay mixture contained 5% glycerol, 10mM Tris.HCl pH 7.5, 10mM EDTA, 1mM DTT, 50mM NaCl, 6.7μM (2μCi) of ^{3}H-AdoMet and between 50ng and 2μg of purified pmt1 protein. Samples were incubated at 28ºC for 1-4h. The sample was then treated with RNase A (10μgml^{-1} for 15 min at 37ºC). This was followed by treatment with proteinase K (100μgml^{-1} for 30 min at 60ºC). The samples were then extracted twice with phenol-chloroform and ethanol precipitated. The samples were resuspended in 20μl TE and spotted onto glass microfibre filters (Whatman). The discs were washed three times in 5% TCA and once in 95% ethanol followed by counting by liquid scintillation. The background was typically 20-30cpm and this was not subtracted from the figures shown in Table I. Each assay was performed three times and the average value was taken. The values for samples containing no substrate were also typically between 20-30 cpm for both cnd1 and M.MspI. The lambda DNA that was used was prepared from *dcm*$^-$ *dam*$^-$ *E. coli* cells and therefore is not methylated at any sites.

iv Preparation of polyclonal antisera against cnd1

1. Histidine-tagged cnd1 protein was purified by nickel-ion chromatography in 500mM NaCl as described above.
2. The 80-1 and 80-2 fractions were dialyzed against buffer A (50mM NaCl) for 3 hour with 3 changes of buffer.
3. The dialyzed samples were centrifuged at 14K in a microcentrifuge for 10 minutes at 4ºC. The pellets were resuspended in buffer C.
4. Three New Zealand white rabbits were each injected with 100μg of cnd1 protein in buffer C which had been mixed with an equal volume of Freund's complete adjuvant in a total volume of 300μl.
5. Four weeks later the rabbits were given a boost of 100μg cnd1 protein

prepared as above but mixed with an equal volume of incomplete Freund's adjuvant in a total volume of 1 ml.

6. Two more boosts were given at 4 week intervals consisting of 100µg of cnd1 protein in 1ml of buffer C.

Ten days after each boost, a 10ml test bleed was taken. Serum and preimmune serum were prepared by standard procedures (Harlow and Lane, 1988) and stored in aliquots at -20ºC.

v Western Blot analysis

Western blotting procedures were carried out using the Bio-Rad mini-trans blot equipment.

1. The cooling unit was filled with water and left at -20ºC to freeze. Fresh transfer buffer (TB) was made up and cooled at 4ºC.

10 X transfer buffer (TB)

Tris base	30.28g
glycine	142.63g
H_2O	to 1 litre

1 X transfer buffer

10 X (TB)	100mls
methanol	200mls
H_2O	700mls

2. Samples were separated by SDS-PAGE and the resulting gel left to equilibrate in TB for 15 minutes at room temperature with agitation.

3. A piece of nitrocellulose (Schleicher and Schuell, BA 83 0.2µM , 401 396) 48mm by 82mm was cut and immersed in H_2O for 5 minutes followed by 10 minutes in TB at room temperature.

4. Two pieces of 3MM paper 70mm by 100mm were cut. These were immersed together with the grey pads from the trans-blot equipment

in TB. The pads were completey wetted.

5. When the gel had equilibrated the apparatus was assembled keeping everything wet by pouring on excess TB as required. Air bubbles were removed from the membrane and 3MM by rolling them with a pipette.

6. Transfer was allowed to proceed for 1 hour at 300mA in the cold room.

7. After transfer the membrane was immersed in blocking solution (5% dried milk powder in TBST) overnight with gentle agitation in the cold room (between 1 and 48 hours). Alternatively the membrane was blotted dry on 3MM paper and stored at room temperature until required. The membrane was re-wetted by immersion in TB for 1-3 minutes, H_2O for 1-3 minutes, TBST for 5 minutes followed by the blocking stage.

   TBST (1 litre) with 0.5%Tween-20: 10ml 1M Tris pH 8.0, 30ml 5M NaCl and 5 ml Tween-20.

Antibody incubations. All washing steps were carried out in TBST (0.5% Tween-20) at room temperature with gentle agitation.

9. The primary incubation was carried out in 5% milk powder in 10ml TBST (0.5% Tween-20) in a petri dish (8.5cm diameter). Incubations were for between 30 and 60 minutes. Polyclonal cnd1 antisera was used at a dilution of 1 in 15000.

10. The filters were washed in a volume of 300ml. After 3 quick rinses the filters were washed once for 15 minutes and then twice for 5 minutes.

11. The secondary antibody incubation was using horseradish peroxidase diluted 1 in 7500 in 25ml TBST (0.5% Tween-20) for 45 minutes.

12. The filters were washed as follows in 300ml as follows: 3 quick rinses, once for 15 minutes, 4 times for 5 minutes each.

13. The filters were developed using the ECL system (Enhanced chemiluminescence, Amersham, RPN 190). 2ml of each ECL solution was mixed in a petri dish (8.5ml diameter) and the filters incubated with gentle agitation for 1 minute. The filters were drained of the developing fluid by pressing their edges against a paper towel. The filters were quickly taped onto a piece of 3MM paper and covered with saranwrap. The filters were exposed to ECL film (Amersham, RPN 2103) for between 15 seconds and five minutes. The films were developed as for other autoradiography procedures.

2.9 *E. coli* manipulations

i Strains

The following *E. coli* strains were used in this work for routine plasmid propagations.

DH5a : *supE44 ΔlacU169 (Φ80 lacZΔMI5) hsdR17 recA1 endA1 gyrA96 thi-1 relA1*

XL1-Blue: *supE44 hsdR17 recA1 endA1 gyrA46 thi relA1 lac⁻ F' [proAB⁺ lacIᵠ lacZΔMI5 Tn10 (tetʳ)]*

The following strain was used for the over-expression of recombinant protein in *E. coli.*

BL21DE3: *pLysS: F⁻ ompT hsdS$_B$ ($r_B^-m_B^-$)*

ii Media and growth conditions

Rich Medium

All *E.coli* strains were routinely grown in rich medium LB consisting of per litre:

Bacto-tryptone	10g
NaCl	10g
Yeast extract	5g

Glucose was added from a sterile stock solution to $1gl^{-1}$ after autoclaving. For solid plate medium $15gl^{-1}$ Bactoagar (Difco, 0140-01) was added before autoclaving.
For large scale preparation of plasmids the cells were grown in an extra rich medium terrific broth.

Bacto-tryptone	12g
Bacto yeast extract	24g
Glycerol	4ml
H_2O	to 1 litre

After autoclaving 100ml of salt solution was added.

KH_2PO_4	2.31g
K_2HPO_4	12.54g
H_2O	to 100ml

Selective antibiotics
Antibiotic stock solutions were stored at -20°C and were added to autoclaved media.

Ampicillin
A 1000x stock solution of $50mgml^{-1}$ of the salt was dissolved in H_2O.

Chloramphenicol
A 1000x stock solution of $34mgml^{-1}$ was dissolved in ethanol.

Tetracycline
A 500x stock of $5mgml^{-1}$ was dissolved in ethanol

Growth conditions
All bacteria were grown at 37°C (except in some of the protein

expression work) in an orbital shaking incubator.

Storage

All bacteria were stored on LB agar plates for up to 4 weeks at 4ºC. Long term storage was in LB medium containing 25%(v/v) glycerol frozen at -70ºC. Bacteria were revived from long term storage by removing a small stab from the frozen culture and streaking out onto selective LB media.

IPTG

Isopropylthio-β-D-galactoside (IPTG) was dissolved in H_2O to give a stock solution of 100mM. This was filtered through a 0.2μM membrane and stored at -20ºC.

iii Plasmid preparation

'Quick preps'

Two methods were used to quickly produce small amounts of DNA (5-10μg).

"Wizard" mini-preps

This method produces very clean DNA suitable for sequencing and transformation into yeast. The Promega wizard kit (A7100) was used according to the manufacturer's instructions.

TELT mini-preps

1. The bacteria were grown in 5-10ml of terrific broth overnight with the appropriate selection.
2. 2-3ml was spun down in an eppendorf.
3. The pellet was resuspended in 800ml TELT.
4. 40μl of 100mgml^{-1} lysozyme in TELT was added and left for 2 minutes.
5. This was boiled for 2 minutes.
6. The tube was cooled for 10 minutes on ice.
7. Cell debris was spun out at 14k in an eppendorf at 4ºC.

8. The supernatant was removed to a fresh tube.
9. 480µl of isopropanol was added.
10. This was left on ice for 2 minutes.
11. The DNA was spun out at 14K for 20-30 minutes at 4ºC.
12. The pellet was washed in 70% ethanol and desiccated.
13. The pellet was resuspended in 100µl TE.

TELT buffer

50mM Tris.HCl pH 7.5
62.5mM EDTA pH 7.5
2.5M LiCl
0.4% Triton X-100

Maxipreps

For large scale preparation of plasmids, the DNA was purified through a caesium chloride gradient.

1. The bacteria were grown in 1 litre of terrific broth with selection.
2. Cells were harvested at 3000rpm for 15 minutes at 4ºC.
3. They were resuspended in 40ml of solution I.
4. 4ml of lysozyme dissolved in solution I ($20mgml^{-1}$) was added and the cells left to incubate for 10 minutes at room temperature.
5. 80ml of freshly made solution II was added and the cells left for 10 minutes on ice.
6. 40ml of cold solution III was added, this was mixed well and the mixture left on ice for 15-30 minutes.
7. The cell debris was spun out at 8000rpm for 5 minutes.
8. The supernatant was filtered through muslin into a graduated cylinder.
9. 0.6 volume of isopropanol was added and mixed well.
10. This was spun at 8000rpm for 10 minutes, the pellet was washed in 70% ethanol and desiccated.
11. The pellet was resuspended in 8ml of TE.
12. 8.8g of caesium chloride was added and the tube inverted several times to dissolve it.
13. 1ml of ethidium bromide ($10mgml^{-1}$) was added.
14. The mixture was spun at 2.5k for 5 minutes.

15. 100ml of Triton-X100 was added to the supernatant.
16. The mixture was syringed into heat sealable tubes, these were balanced to within 0.1g, sealed and spun in an NVT90 rotor on the following programme.

90K	90 minutes
87K	15 minutes
83K	15 minutes
81K	30 minutes
80K	30 minutes

17. After centrifugation the plasmid band was removed with a sterile syringe.
18. The ethidium bromide was extracted with an equal volume of isoamyl alcohol (this was done 4-5 times until the DNA solution became clear. Each time the mixture was spun for 3 minutes at 1.5K to separate the phases).
19. 3 volumes of 70% ethanol were added and the DNA precipitated at 10000rpm for 20 minutes.
20. The pellet was washed, dried and resuspended in 2ml of TE.
21. 1/10th volume of sodium acetate and 2 volumes of ethanol were added.
22. This was spun as in 19, washed and dried and the DNA resuspended in TE (between 100µl and 500µl).

Solution I
50mM glucose
25mM Tris.HCl pH 8.0
10mM EDTA pH 8.0

Solution II
0.2M NaOH
1% SDS

Solution III
3M KOAc
2M acetic acid

iv Transformation

Preparation of cells

1. 1 litre of L-broth was inoculated with 1/100th volume of a fresh overnight culture.
2. The cells were grown at 37°C with vigorous shaking to an OD_{600} of between 0.5 and 1.0.
3. The flask was chilled in iced water for 30 minutes and the cells were then spun down in JB-6 centrifuge for 15 minutes at 3000rpm at 4°C.
4. The pellet was resuspended very gently in 1 litre of ice cold water.
5. The cells were centrifuged as in step 3.
6. The cells were resuspended in 1 litre of cold water.
7. The cells were centrifuged as in step 3.
8. The cells were resuspended in 250ml of cold water and spun in a JA-14 rotor at 5000rpm for 10 minutes.
9. The pellet was resuspended in 40ml of ice cold 10% glycerol. The OD_{600} was measured.
10. The cells were resuspended in a final volume of 2-3ml in 10% glycerol to give a concentration of approximately 3×10^{10} cells ml^{-1}.
11. The cell suspension was divided into 40µl aliquots, frozen on dry ice and stored at -70°C. The cells are good for at least 6 months under these conditions.

Electro-transformation

1. The cells were thawed on ice.
2. In a cold 1.5ml eppendorf tube, 40µl of the cell suspension was mixed 1-2µl of DNA (the DNA should be in a low ionic strength buffer such as TE). The cells were left on ice for 1 minute.
3. The gene pulser apparatus was set at 25mF and 2.5kV (for DH5-α cells, 1.8kV was used). The pulse controller was set to 200Ω.
4. The mixture of cells and DNA was transferred to a chilled electroporation cuvette, and the suspension was shaken to the bottom so that the cells made contact with both sides. The cuvette was placed in the safety chamber slide and the slide pushed into the chamber

until the cuvette was sealed between the contacts in the base of the chamber.

5. The cells were pulsed once.
6. The cuvette was removed from the chamber and immediately 1ml of [1]SOC medium was added to the cells. A pasteur pipette was used to resuspend the cells. (This rapid addition of SOC after the pulse is very important in maximizing the recovery of the transformants).
7. The cell suspension was transferred to a 50ml falcon tube and left to shake for 1 hour at 37ºC to allow the cells to recover.
8. 100ml of cells was plated on a selective medium.

[1]SOC buffer
2% bacto tryptone
0.5% bacto yeast extract
10mM NaCl
2.5mM KCl
10mM $MgCl_2$
20mM glucose

2.10 Fission yeast manipulations

i Strains

The wild type strain and mutant strains of fission yeast *Schizosaccharomyces pombe Linder*, were all derived from the haploid heterothallic *972 h⁻* and *975 h⁺* isolates described by Leupold (1950). The details of the strains used in this project are given below.

strain
975 h+
972 h-
cnd1-1 (renamed pat1-8) leu1-32 h⁻
pat1-114 leu1-32 h⁻
ade6-M210 ura4-D18 leu1-32 h-
ade6-M216 ura4-D18 leu1-32 h+
leu1-32 ade6-M210 ura4D/SE Ch16 ade6-M216 M23::LEU2 h⁻ (a gift from

Robin Allshire)
pmt1::ura4 leu1-32 ade6-M210 h+
pmt1::ura4 leu1-32 ade6-M210 h90
pmt1::ura4 leu1-32 ade6-M210 ura4-D18 h+
pmt1::ura4 leu1-32 ade6-M210 ura4-D18 Ch16 ade6-M216 M23::LEU2 h+

ii Media and growth conditions

Media

Strains of fission yeast were routinely grown on solid YE (Difco, 0127-01-7) complex medium supplemented with adenine and uracil.
YE contains per litre:

glucose	30g
yeast extract	5g
adenine	75mg
uracil	75mg
Difco-Bacto agar	15g

For complex liquid medium YE was used minus Difco-Bacto agar.
EMM (Moreno *et al.*, 1991) was used as minimal medium, and is a modification of EMM 2 (Mitchison, 1970 and Nurse, 1975) and contains the following per litre:

glucose	20g
KH phthalate	3g
Na_2HPO_4	1.8g
NH_4Cl	5g
$NaSO_4$	100mg
$CaCl_2$	15mg
$MgCl_2$	1g
vitamins	1ml
minerals	100ml

Where vitamins are : inositol, 5g; nicotinic acid, 5g; calcium pantathenate, 0.5g; and biotin, 5mg all dissolved in 500ml H_2O.
Where minerals are H_3BO_3 1g; $MnSO_4.4H_2O$, 1.04g; $ZnSO_4.7H_2O$, 800mg; $FeCl_3.6H_2O$, 400mg; H_2MoO_4, 288mg; $CuSO_4.5H_2O$ 80mg; Citric acid, 2g; and KI, 20mg, all dissolved in 200ml H_2O. Both were stored at 4°C.

Growth supplements such as adenine and uracil were made as 50 X stock solutions of 375mgml^{-1} and leucine was made up as 750mgml^{-1}. All supplements were added as required after autoclaving. Thiamine was made as a 100mM solution and stored in the dark. It was used at a final concentration of 4μM.

For the induction of conjugation and meiosis, nitrogen limiting medium ME, consisting of 30gl^{-1} malt extract and 20gl^{-1} agar was used.

To check ploidy of cells, phloxin B was added from a stock solution of 5gl^{-1} to a final concentration of 2.5mgl^{-1} in solid media. Diploid cells die faster than haploid cells and accumulate the dye more quickly so they can be differentiated by relative colour staining. To check the mating type of a particular strain a loopful of cells were placed on a malt extract plate with a loopful of a h^+ strain and on another part of the plate with a loopful of a strain known to be h^-. The strains were mixed with a loopful of sterile H_2O. After three days incubation at 28ºC the patches were exposed to iodine crystals for 1-3 minutes. If mating had occurred between two strains of opposite mating types then the patch would stain grey-black. This was due to the starch in the spores reacting with the iodine vapour.

Storage

Short term storage of fission yeast (up to 4 weeks) was on solid medium (but not containig phloxin B) at 4ºC. Long term storage was in medium containing 20% (v/v) glycerol at 70ºC.

iii Genetical analysis

Crossing strains

The standard genetical procedures of Gutz *et al.*, (1974) and Kohli *et al.*, (1977) were followed. Strains were crossed by mixing together, with a loopful of sterile water, fresh isolates of two cell types (h^+ and h^-) on the surface of an ME plate. The mating mix was incubated at 28ºC for 2-3 days to allow zygotes to form. The progeny of the crosses were examined by either tetrad or random spore analysis.

Tetrad analysis

A loopful of a 2-day old cross was placed in a line on a YE plate. Using a micromanipulator, asci were placed 5mm apart and incubated overnight at 20ºC to allow the asci walls to break down. Each ascus was then micromanipulated to give a line of 4 separate spores. The spores were incubated for 3-5 days and then replica plated for further analysis. With this technique the products of a single meiosis can be analyzed and unlike random spore analysis, non-viable meiotic classes of spores can be identified.

Random spore analysis

A loopful of 3-day old mating mix was suspended in 500µl of sterile distilled water containing 10µl of a stock solution of the snail gut enzyme (Suc d'helix pomatia, Industrie Biologique, France) and incubated overnight at 28ºC. The stock solution is a 1/10th dilution of the preparation supplied, which is stored at 4ºC in the dark. The spore concentration was estimated using a haemocytometer and an appropriate dilution, to give about 500 spores per plate, spread onto solid YE media.

Analysis of phenotypes

Phenotypes of cells were tested by replica plating or by streaking from a master plate onto EMM plus or minus growth supplements for auxotrophs and onto YE with phloxin at the restrictive and permissive temperatures for temperature-sensitive strains.

Diploid construction

For a sporulation competent strain, h^+/h^- strains were constructed using complementing alleles of *ade6*. The *ade6-M210* and *ade6-M216* alleles both confer upon the cells, a requirement of adenine for growth, but heterozygous diploids *ade6-M210 ade6-M216* are prototrophic. The advantage of this system is that very little recombination occurs between these two loci, so that very few prototrophic haploids arise.

Strains carrying the *ade* mutated alleles were crossed in the normal way, left to conjugate overnight at 28ºC and then streaked onto selective media for adenine prototrophs. Diploid colonies were recognized by their colour on phloxin-containing plates, and their ability to sporulate which was

checked microscopically.

Stability test

The procedure described below is useful for distinguishing between an autonomously replicating plasmid and an integration event. The rationale of the procedure is that if a plasmid replicates autonomously, it will be lost in the absence of selection. On the other hand, if the plasmid has integrated, or if there has been a reversion or gene conversion event, the phenotype will be maintained after relaxing the selection.

1. Transformants were replica-plated onto selective media to reduce background growth.
2. The transformants were then plated onto non-selective media (YE).
3. These patches were then replica-plated onto non-selective media three times successively.
4. The transformants were replica-plated onto selective media.
5. Transformants were then streaked to single colonies onto non-selective media (YE).
6. The colonies were replica-plated to selective media.
7. Transformants from which all colonies grew up were considered to be stable.

iv Cell physiology

For all experiments examining aspects of fission yeast cell physiology, actively growing cells were used. Such cells were undergoing exponential growth, and were at a concentration of 3-8 x 10^6 cells ml^{-1}. Exponentially growing cells were obtained by inoculating a single colony into a 10ml YE preculture and incubating until the cells had entered stationary phase of growth (1-2 days). An aliquot of the preculture(1/500-1/100th) was inoculated into 200ml of EMM or YE in a 500ml conical flask, and incubated in an orbital shaker for 14-18 hours at 25-32ºC. Cell number was monitored until cells reached mid-exponential stage.

Estimation of cell number

Cell number was estimated using a coulter counter (Industrial D). A 100µl sample of cell culture was removed and mixed in 10ml of ISOTON solution (a 1/1000 dilution), sonicated for 5 seconds at setting 2 on a Lucas Soniphore sonicator. Usually 2 counts of 500µl were taken and combined to give a cell count per µl.

v Transformation

Electroporation using Bio-Rad equipment was used to transform in both plasmids and linear fragments of DNA into fission yeast cells.

1. Cells were grown to an OD_{595} of between 0.25 and 0.5 (5×10^6 - 1×10^7 cells ml^{-1}).
2. The cells were harvested at 5K in a JA-14 rotor at room temperature for 10 minutes.
3. The pellet was resuspended in 20ml of ice cold 1.2M sorbitol and transferred to a 50ml falcon tube.
4. The cells were harvested at 2.5K for 5 minutes in a benchtop centrifuge.
5. The pellet was resuspended in 10ml of cold sorbitol and spun as in step 4.
6. Steps 4 and 5 were repeated.
7. After centrifugation (again as in step 4) the pellet was resuspended in 1.2M sorbitol to give a concentration of 1×10^9 cells ml^{-1}.
8. The cells were divided into 200µl aliquots and mixed with between 50ng and 10µg of DNA.
9. After mixing with DNA the cells were immediately pulsed, using settings of 2.25kv, 200Ω and 25mF.
10. After the pulse, 1ml of ice cold 1.2M sorbitol was added.
11. 100-500µl of cells were spread on selective medium and left to incubate at 25-32°C.
12. Transformants appeared after 2-5 days.

vi Preparation of nucleic acids from fission yeast

Genomic DNA

1. 10ml cultures of media were inoculated with a single yeast colony and incubated until the culture reached stationary phase (1-2 days).
2. Cells were harvested (5 minutes at 2500rpm) in 15ml centrifuge tubes.
3. Supernatants were discarded and cells resuspended in 1.5ml [1]CPS containing 2.5mgml^{-1} zymolyase 20T (Seikaguku Kogyo 120491): suspensions were transferred to eppendorf tubes.
4. Cells were incubated for 1 hour at 37ºC.
5. Cells were harvested by centrifugation at 14K for 2 minutes.
6. Supernatants were discarded and cells resuspended in 300µl of 5 x TE; 35µl of 10% SDS (w/v) was added and the mixture incubated at 65ºC for 5 minutes.
7. 100µl of 5M potassium acetate (pH 5.6) was added and the mixture incubated on ice for 30 minutes.
8. Cells were centrifuged at 4ºC at high speed for 15 minutes: the supernatant removed (400µl) and to this 1ml of ice-cold 95%(v/v) ethanol added and mixed.
9. This was centrifuged at high speed for 10 minutes at 4ºC.
10. The supernatant was discarded and the pellet resuspended in 400µl of 5 x TE containing 100mgml^{-1} RNAse (bovine pancreatic Sigma R - 5000): incubated for 2-4 hours at 37ºC.
11. The suspension was extracted once each with an equal volume of phenol, phenol-chloroform and chloroform.
12. The final aqueous phase was transferred to another eppendorf tube, sodium acetate precipitated and washed with 70% ethanol.
13. Pellets were dried, dissolved in 100µl of TE and DNA quality and amount checked.

[1]CPS

50mM citrate-phosphate buffer pH 5.6
1.2M sorbitol
0.1% (v/v) β-mercaptoethanol (added immediately before use)

Total RNA

This method based on that of Kaufer *et al.*, (1985) produced fission yeast total RNA which was used for Northern blotting. Gloves were worn during all RNA procedures to prevent RNAse contamination.

1. 200 ml cultures of fission yeast cells in mid exponential growth were harvested by centrifugation at 5000rpm for 10 minutes in four 50ml falcon tubes (Falcon 7020).
2. Each pellet was resuspended in 1ml of [1]sterile saline and transferred to an eppendorf tube; cells were pelleted by centrifugation at high speed, and the supernatant discarded (cells could be stored at -70ºC at this stage).
3. Cells were resuspended in 75µl of [2]STE.
4. Hydrochloric acid washed beads (425-600 micron, Sigma G-9268) were added to just beneath the meniscus, and the cells briefly vortexed.
5. 600µl of [3]NTES was added and the mixture briefly vortexed again.
6. 500µl of hot phenol (65ºC) was added.
7. The mixture was incubated at 65 ºC for 5 minutes with frequent vortexing.
8. The mixture was centrifuged for 1 minute at high speed and the aqueous phase and protein interface were removed to a second 500µl aliquot of hot phenol.
9. This was incubated again at 65ºC for 2 minutes with frequent vortexing.
10. This was centrifuged for 1 minute at high speed and again the aqueous phase and interface were removed to a third 400µl aliquot of hot phenol.
11. This was incubated at 65ºC for 2 minutes with frequent vortexing.
12. The aqueous phase only was removed into 400µl of phenol-chloroform at room temperature, vortexed and spun at high speed for 1 minute.
13. The aqueous phase was re-extracted with 300µl of chloroform, vortexed and spun for 1 minute at high speed.
14. The aqueous phase was ethanol precipitated.
15. The supernatant was discarded and RNA pellet washed with 70%

ethanol in DEPC.H_2O.

16. The RNA pellet was resuspended in 55µl of ice-cold DEPC.H_2O
17. 5µl of each sample was used to estimate the amount and quality by spectrophotometry.
18. The RNA was stored at -70°C.

[1]sterile saline
0.9% (w/v) NaCl

[2]STE
0.32M sucrose
20mM Tris.HCl (pH 7.5)
10mM EDTA (pH 8.0)
0.5mgml^{-1} heparin (Sigma, H-7005) added solid just prior to use

[3]NTES
100mM NaCl
5mM EDTA
50mM Tris.HCl (pH 7.5)
1% (w/v) SDS
0.5mgml^{-1} heparin added solid just prior to use

vii Preparation of protein from fission yeast

This method produces total non-denatured fission yeast protein extracts which were used for western blots.

1. 200ml cultures of fission yeast cells in mid exponential phase of growth were harvested by centrifugation at 2500 rpm for 5 minutes in four 50ml falcon tubes.
2. Supernatants were discarded and each cell pellet resuspended in 100µl of [1]lysis buffer in eppendorf tubes: acid washed glass beads (425-600 micron, Sigma G-9268) were added to just beneath the meniscus.
3. Cells were broken by vigorous vortexing: 15 minutes of 30 seconds with 30 second intervals on ice.

4. Cell walls were pelleted by centrifugation at high speeds for 5 minutes at 4ºC.
5. Supernatants were transferred to another eppendorf tube, and the protein extract clarified by centrifugation at 16000rpm (Beckman JA-18) for 30 minutes at 4ºC.
6. Supernatants were transferred to another eppendorf tube and snap frozen on solid CO_2 and stored at -70ºC.

[1]Lysis buffer

	stock	
50mM KCl	3M	17µl
50mM Tris.HCl pH 7.9	1M	50µl
25% glycerol	50%	500µl
2mM DTT	1M	2µl
0.1% Triton X-100	1%	100µl
chymostatin (Sigma, C 7268)	5mgml^{-1}	5µl
antipain (Sigma, A-6271)	5mgml^{-1}	5µl
leupeptin (Sigma, L-2884)	5mgml^{-1}	5µl
pepstatin (Sigma, P-4265)	5mgml^{-1}	5µl
H_2O		303µl
[a]0.2mM PMSF	100mM	2µl

DTT: dithiothreitol (Sigma D-0632) 1M solution in H_20, stored at -20ºC.
PMSF: phenylmethanesulphonylfluoride (Sigma, P-7626) 100mM solution in isopropanol, stored at 4ºC. [a]added to the buffer just before use.
Protease inhibitors: each at 5mgml^{-1}, stored at -20ºC

The protein concentration was assayed using the Bio-Rad protein assay kit (500-0002) according to the manufacturer's instructions.

Chapter 3 DNA methylation analysis of the *S. pombe* genome

3.1 Introduction

Previous attempts to detect DNA methylation in *S. pombe* consisted of analyses of the whole genome by biochemical methods and a visual comparison of genomic digests using methylation-sensitive and methylation-insensitive restriction enzymes (Antequera *et al.*, 1984). However, if DNA methylation is localized to specific areas of the *S. pombe* genome, it may be difficult to detect against the background of the genome as a whole. One way to extend the search for m^5C would be to analyze the methylation status of a particular region of the *S. pombe* genome. This analysis could be carried out by hybridizing specific probes to genomic DNA digested with restriction enzymes which have different sensitivities to methylated DNA, but which otherwise recognize the same target sequence.

Stimulated by the identification of the *cnd1* m^5C-MTase homologue in fission yeast, an analysis of this nature was carried out concentrating on regions that contain repetitive DNA or sequences that are themselves highly repeated within the genome. In addition, nearest neighbour analysis was performed to assess whether CNG methylation occurs in the fission yeast genome. Finally, *S. pombe* DNA was fractionated over a methylated DNA binding column in an attempt to identify methylated sequences in the *S. pombe* genome. The results of these studies are presented in this chapter.

3.2 Are CNG sequences methylated in *S. pombe*?

Before carrying out Southern blot analyses of specific genomic regions, *S. pombe* DNA was digested with isoschizomeric pairs of restriction enzymes including *HpaII/MspI* and *BstNI/EcoRII*. *HpaII* and *MspI* both recognize the sequence CCGG, but only the latter enzyme will cut if the internal cytosine is methylated. *BstNI* and *EcoRII* both recognize the sequence CCWGG (where W = A or T), however, the methylated sequence $C^{m5}CWGG$ is resistant to *EcoRII* cleavage while remaining susceptible to *BstNI* digestion (Figure 3.1).

Fission yeast DNA was digested with these enzymes and the resulting restriction patterns analyzed by electrophoresis (Figure 3.2). As expected from the results described by Antequera *et al.*, (1984), no differences could be identified between the patterns generated by *HpaII* and *MspI*.

Surprisingly, there was a substantial difference between the restriction patterns of *BstNI* and *EcoRII* digested DNA. A significant proportion of the CCWGG sites in the *S. pombe* DNA appeared to be uncut by *EcoRII* suggesting that there was methylation of the internal cytosine in CCWGG sequences. This result was reproducible even after re-digestion of the DNA following phenol-chloroform extraction and ethanol precipitation (data not shown). The extent of CNG methylation implied by this result was surprising given the failure to detect methyl groups at any cytosine in the genome as a whole (Antequera *et al.*, 1984).

The properties of *EcoRII*, however, suggested an alternative explanation for its inability to cleave *S. pombe* DNA. It has been reported that *EcoRII* requires the presence of at least two recognition sites in the target DNA for activity (Kruger *et al.*, 1988). Moreover, DNA with several

A

	Cleavage by enzymes	
recognition sequence	*MspI*	*HpaII*
CCGG	yes	yes
$C^{m5}CGG$	yes	no

B

	Cleavage by enzymes	
recognition sequence	*BstNI*	*EcoRII*
CCWGG	yes	yes
$C^{m5}CWGG$	yes	no

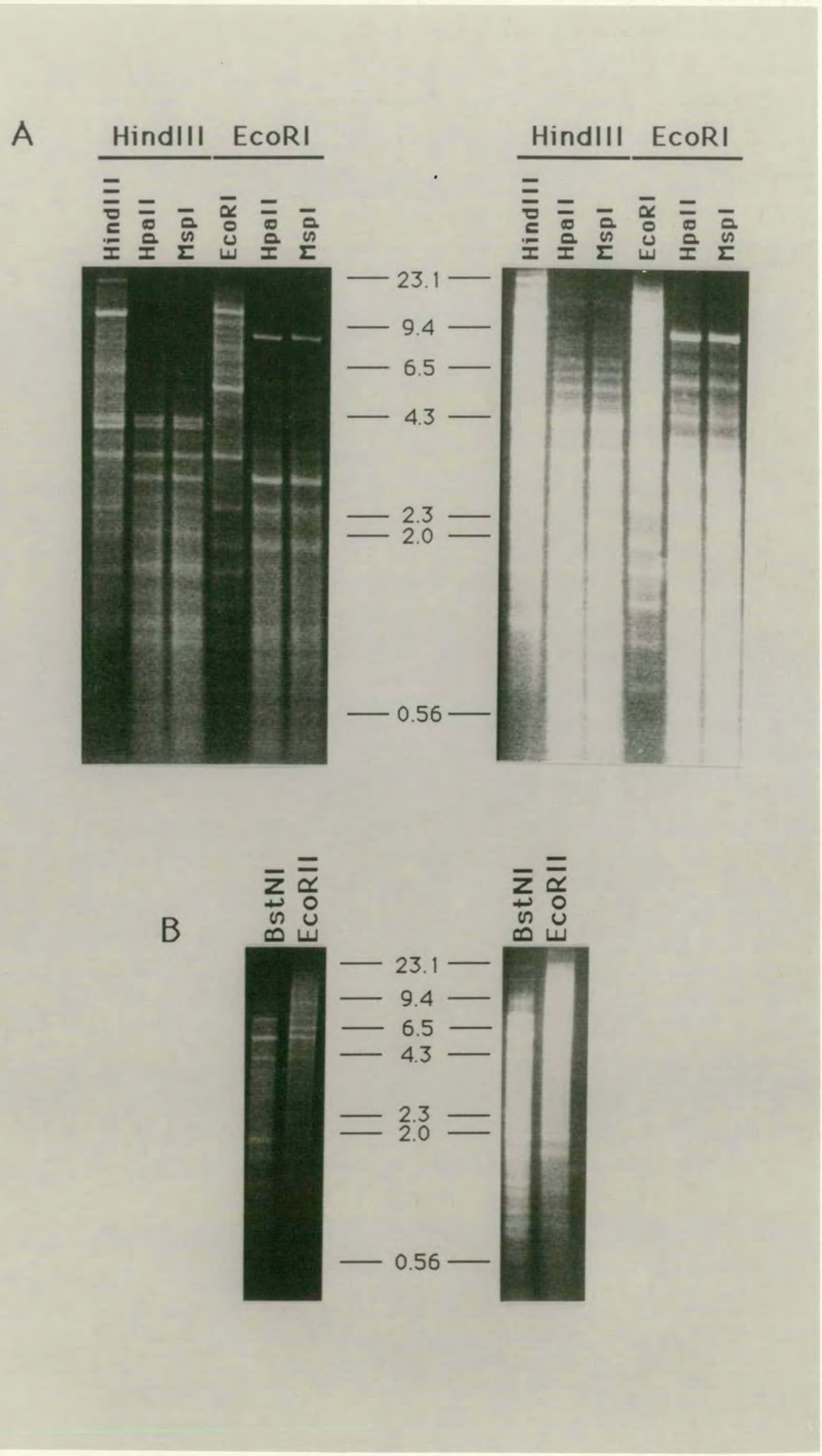
Figure 3.1 Isoschizomeric pairs of restriction enzymes used to analyze the methylation status of the *S. pombe* genome.

Both *HpaII* and *MspI* cleave the DNA in between the two cytosine residues. *EcoRII* cleaves the DNA 5' to the external cytosine whereas *BstNI* cleaves in between the internal cytosine and the middle residue of the 5bp sequence. This can be either adenine or thymine (shown as W in the figure).

Figure 3.2 Digestion of *S. pombe* genomic DNA using methylation sensitive restriction enzymes

A *S. pombe* genomic DNA (1µg) was digested with *HindIII* or *EcoRI* overnight at 37ºC. After extraction with phenol/chloroform and ethanol precipitation, the DNA was digested with either *HpaII* or *MspI* as indicated. DNA was separated on a 0.8% agarose gel. Molecular weight standards are indicated in kilobases. Two different exposures of the same gel are shown so that specific fragments as well as the overall pattern of digestion can be identified.

B *S. pombe* genomic DNA (1µg) was digested with *BstNI* or *EcoRII* overnight at either 60ºC or 37ºC respectively. DNA was separated on a 0.8% agarose gel. Molecular weight standards are indicated in kilobases. Two different exposures of the same gel are shown.


A
HindIII
EcoRI
HindIII
EcoRI
HindIII
HpaII
MspI
EcoRI
HpaII
MspI
HindIII
HpaII
MspI
EcoRI
HpaII
MspI
23.1
9.4
6.5
4.3
2.3
2.0
0.56
B
BstNI
EcoRII
BstNI
EcoRII
23.1
9.4
6.5
4.3
2.3
2.0
0.56

recognition sites is only cleaved when the distance between them does not exceed the limit of about 1000 base pairs, the efficiency of cleavage decreasing with inter-site distance (Pein *et al.*, 1991). It seemed possible therefore, that the inability of *EcoRII* to cut DNA to the same extent as *BstNI* was attributable to the distance between sites rather than their methylation status. To resolve this issue, nearest neighbour analysis of *S. pombe* DNA cut with *BstNI* was carried out (Figure 3.3A). This method relies on the positive identification of a signal corresponding to m^5C rather than inference from a lack of digestion.

Genomic DNA from *S. pombe* was cut with *BstNI* and the resulting DNA termini filled in. The DNA was then digested to give deoxynucleoside 3' monophosphates which were separated by two dimensional chromatography and the resulting signals identified by autoradiography (Figure 3.3B). There was no signal corresponding to 5-methylcytidine. As shown in Figure 3.3A, the expected position of 5-methylcytidine is close to cytidine but even after prolonged exposures there were no additional signals. Slight nicking of the genomic DNA during isolation leads to signals of lower intensity corresponding to thymidine, adenosine and guanosine.

The resolution of nearest neighbour analysis is such that 5-methylcytosine will only be detected when present at levels of greater than 0.1% of the total cytosine (Gruenbaum *et al.*, 1981b). Therefore, these results clearly contradicted the results of the restriction enzyme analysis. The *EcoRII/BstNI* isoschizomeric restriction enzyme analysis detects m^5C by the failure of *EcoRII* to cut DNA. As described above, there are other reasons why digestion by *EcoRII* might not be complete. In contrast, nearest neighbour analysis relies on the direct chemical detection of m^5C. With this technique, the absence of a signal corresponding to m^5C can only be due to

the lack of a significant amount of this base in the sample being analyzed. It was concluded, therefore, that the *EcoRII* sites are not extensively methylated in *S. pombe* DNA and thus in turn, it is unlikely that the fission yeast genome contains widespread CNG methylation.

3.3 Methylation analysis of repetitive regions in the *S. pombe* genome using the isoschizomeric pair of restriction enzymes *HpaII* and *MspI*

To assess the methylation status of specific regions of the *S. pombe* genome, Southern blots of DNA digested with *HpaII* and *MspI* were hybridized with probes made from specific regions of repetitive DNA (Figure 3.4). In most cases, the DNA was also digested with either *HindIII* or *EcoRI* to reduce fragment size and allow resolution of potential small size differences.

The first probe that was used consisted of the 10.4kb *HindIII* fragment containing the *S. pombe* genes for the 18S, 5.8S and the 28S ribosomal RNAs (Barnitz *et al.*, 1982). This unit is highly repeated in the *S. pombe* genome and rDNA genes have been found to be methylated in other organisms (see Introduction). No differences were detected between the *HpaII* and *MspI* patterns when hybridized to the rDNA probe indicating that the internal cytosine within CCGG sequences in this region are not methylated (Figure 3.5).

In order to assess methylation in repeated sequences in general, total *S. pombe* genomic DNA was radio-labelled and used as a probe in a Southern hybridization of *S. pombe* DNA digested with *HpaII* and *MspI*. The rationale behind the experiment was that if a small amount of genomic DNA was radioactively labelled, only regions that are highly repeated will give a

Figure 3.3A Nearest neighbour analysis of DNA digested with *BstNI*

DNA is cut with *BstNI* which cleaves CCWGG sequences (where W is A or T) between the internal cytosine and the A or T residue. The resulting termini are filled in with the Klenow fragment of DNA polymerase I using ^{32}P-labelled dATP resulting in a radioactively-labelled phosphate being incorporated adjacent to the internal cytosine of the *BstNI* site. The DNA is then treated with micrococcal nuclease and spleen phosphodiesterase to give deoxynucleoside 3' monophosphates. This digestion results in the radioactively-labelled phosphate being effectively transferred from the adenosine moiety to the internal cytosine of the *BstNI* site. The resulting deoxynucleoside 3' monophosphates are then separated by 2-dimensional thin layer chromatography and the radioactively-labelled bases detected by autoradiography. The expected location of 5-methylcytidine relative to that of cytidine is shown. p = radio-labelled phosphate. 1 and 2 indicate the directions of the thin layer chromatography.

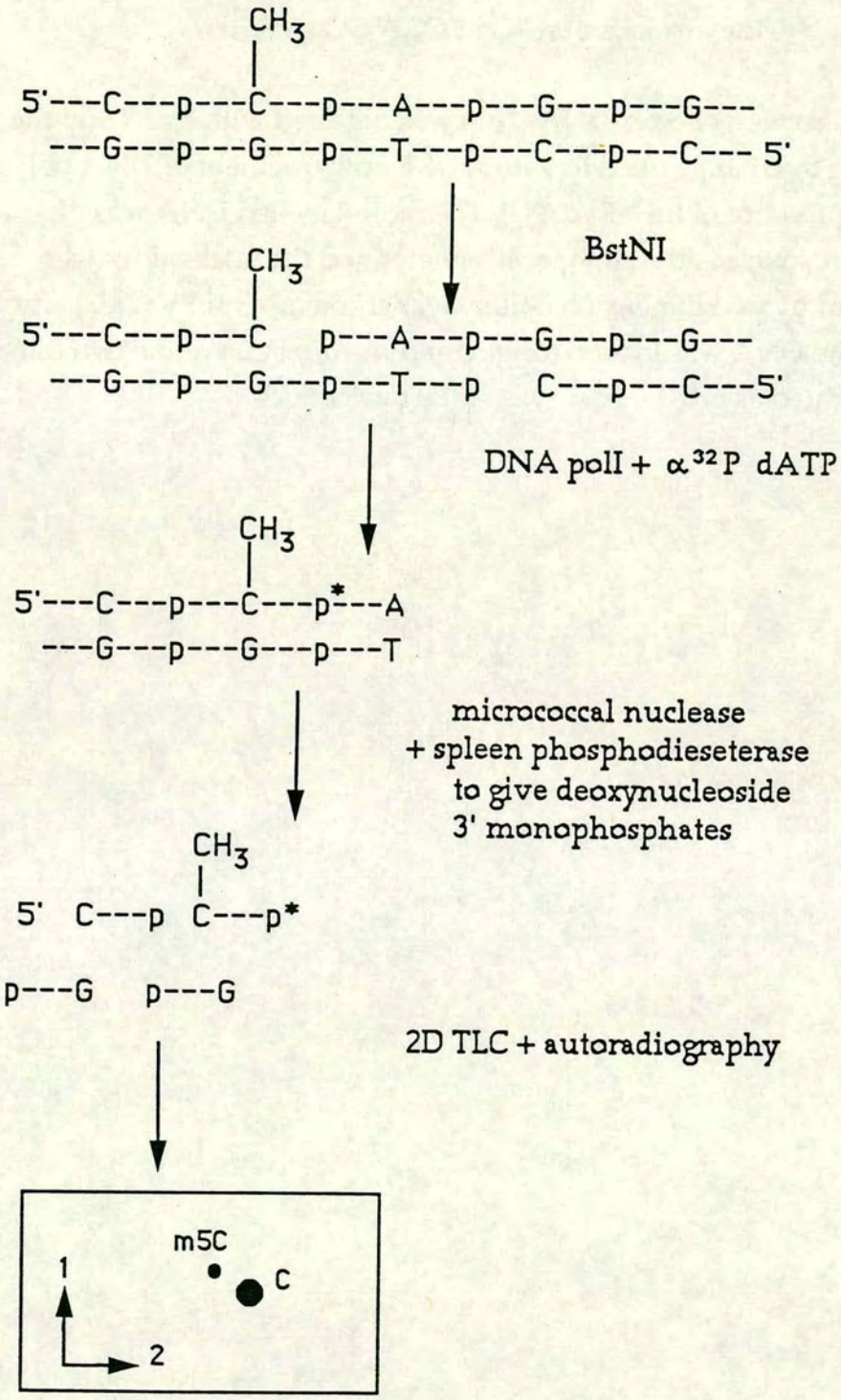
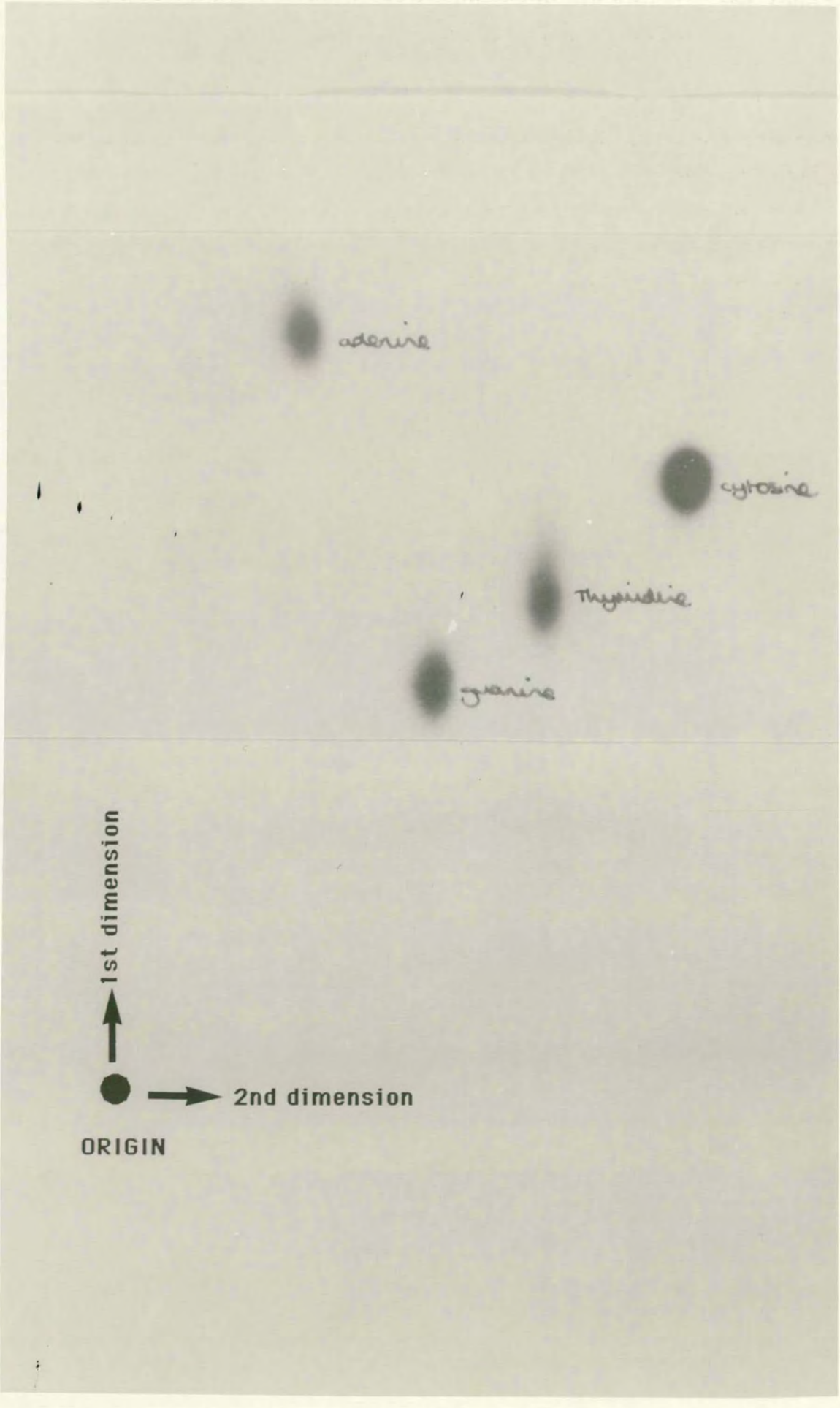


Figure 3.3A Nearest neighbour analysis of DNA digested with BstNI

Figure 3.3B Nearest neighbour analysis of *S. pombe* DNA illustrating the absence of m^5C in CCWGG sites

S. pombe genomic DNA (2μg) was digested with *BstNI* and the resulting overhangs filled in with the Klenow fragment of DNA polymerase I in the presence of [α-^{32}P] dATP. The radio-labelled DNA was digested to give deoxynucleoside 3'-monophosphates and the nucleotides were separated by two dimensional thin layer chromatography (TLC). The point where the DNA was loaded (origin) and the direction of the two dimensions are indicated as are the positions of the nucleotides.


adenine
cytosine
Thymidine
guanine
1st dimension
2nd dimension
ORIGIN

significant signal upon hybridization. The majority of bands that are seen in such a blot will represent mitochondrial DNA as this is the most repeated DNA in the *S. pombe* genome (Anziano *et al.*, 1983). The total genome probe hybridized to the same bands in the *HpaII* and *MspI* lanes (Figure 3.6A), again indicating a lack of methylation at these sites. Slight differences in intensity of some of the bands was noted, but this was not reproducible and therefore not likely to have been caused by methylation.

A similar result was obtained using the long terminal repeat (LTR) which is present in the two families of transposable elements that have been identified in *S. pombe* (Levin *et al.*, 1990). A probe containing the LTR was generated by PCR from *S. pombe* genomic DNA (see Appendix A). It was not possible to detect any methylated CCGG sites in the regions detected by the LTR probe.

The next sequences to be assessed were the *S. pombe* centromeres. They contain many copies of two repeats known as dg and dh (Murakami *et al.*, 1991). It seemed reasonable to propose that these regions might be methylated because in addition to their repetitive nature, abnormal centromere function due to hypomethylation could readily explain the defective nuclear division seen in the *cnd1-1* mutant. For the same reasons, the methylation status of the *S. pombe* telomeres was also examined. As shown in Figure 3.7, no differences were found between *HpaII* and *MspI* digests in regions hybridizing to either the centromeric or telomeric probes.

3.4 Fractionation of *S. pombe* DNA using a methylated DNA binding column

As the search for DNA methylation in specific regions of the *S. pombe* genome had not proved successful, a different approach was taken to try to

Figure 3.4 Probes used in the methylation analysis of the *S. pombe* genome

A Restriction map of the *S. pombe* rDNA repeat (Barnitz *et al.*, 1982). The entire 10.4kb *HindIII* fragment was used as a probe. The location of the 18S, 5.8S and 28S ribosomal RNA genes are represented by black boxes.

B Restriction map of the *S. pombe* Tf1-107 transposable element (Levin *et al.*, 1990). The arrows represent the 358bp LTR which was used as a probe. The black boxes represent the functional domains identified on the basis of the amino acid similarity to the protease (PR), reverse transcriptase (RT), RNase H (RH) and integrase (IN) domains of retroviruses.

C Restriction map of the *S. pombe* centromeric dg and dh repeats from centromere III (Murakami *et al.*, 1991). Arrowheads indicate the direction of the repeat elements dg and dh (found in all three *S. pombe* centromeres). The 6.1kb *EcoRI* fragment was used as a probe.

D Restriction map of the *S. pombe* telomeric regions. The diagram shows the telomere repeats and telomere-associated sequences which are represented by open and black boxes respectively (Sugawara and Szostak, 1986). The 7.1kb *HindIII-EcoRI* fragment was used as a probe.

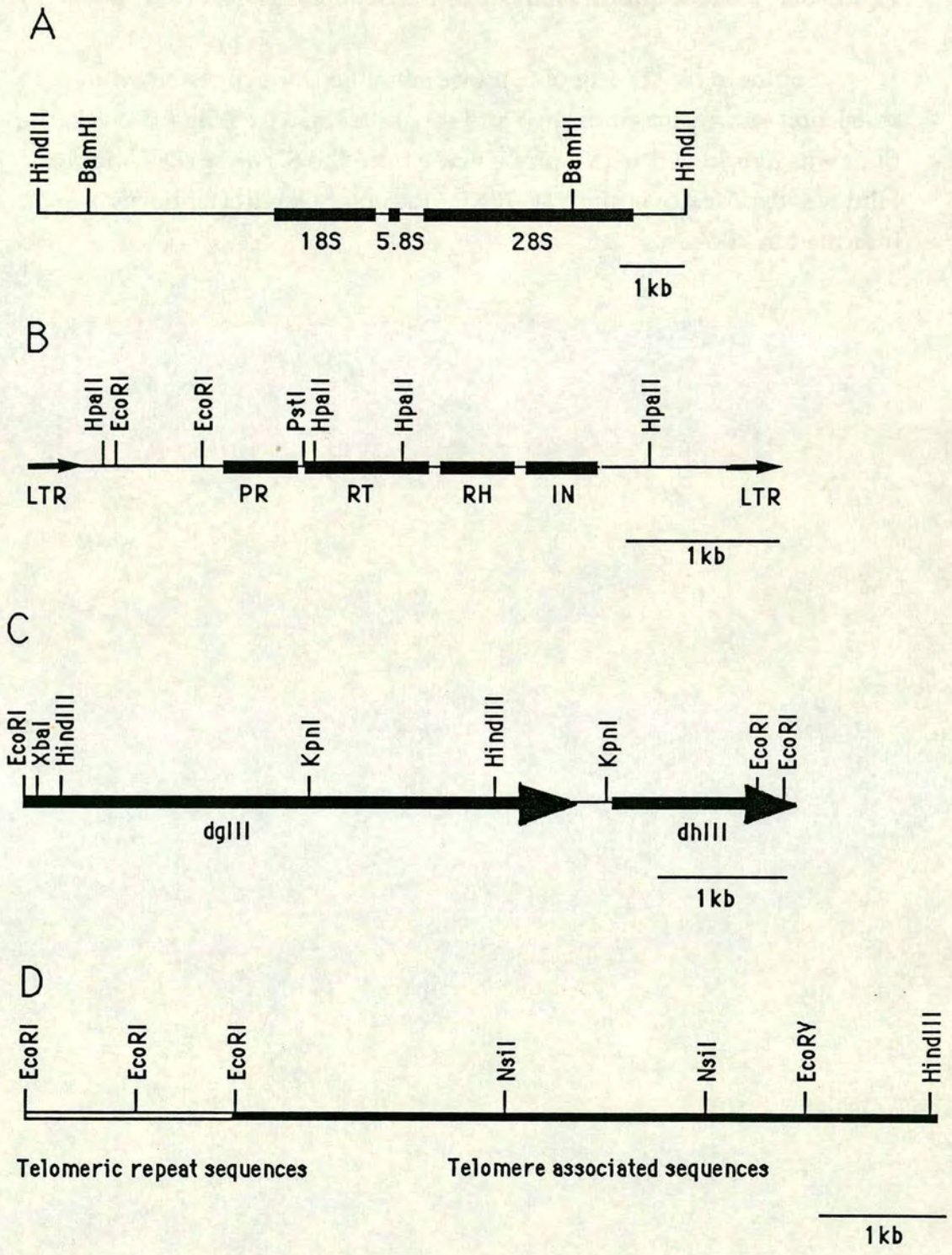


Figure 3.4 Probes used in the methylation analysis of the *S. pombe* genome

Figure 3.5 Methylation analysis of the *S. pombe* rDNA repeats

Southern blot of 1μg of *S. pombe* genomic DNA digested with restriction enzymes as indicated and separated on a 0.8% agarose gel. The filter was hybridized with a probe made from the *S. pombe* rDNA repeat. Film was exposed overnight at -70°C. Molecular weight standards are indicated in kilobases.

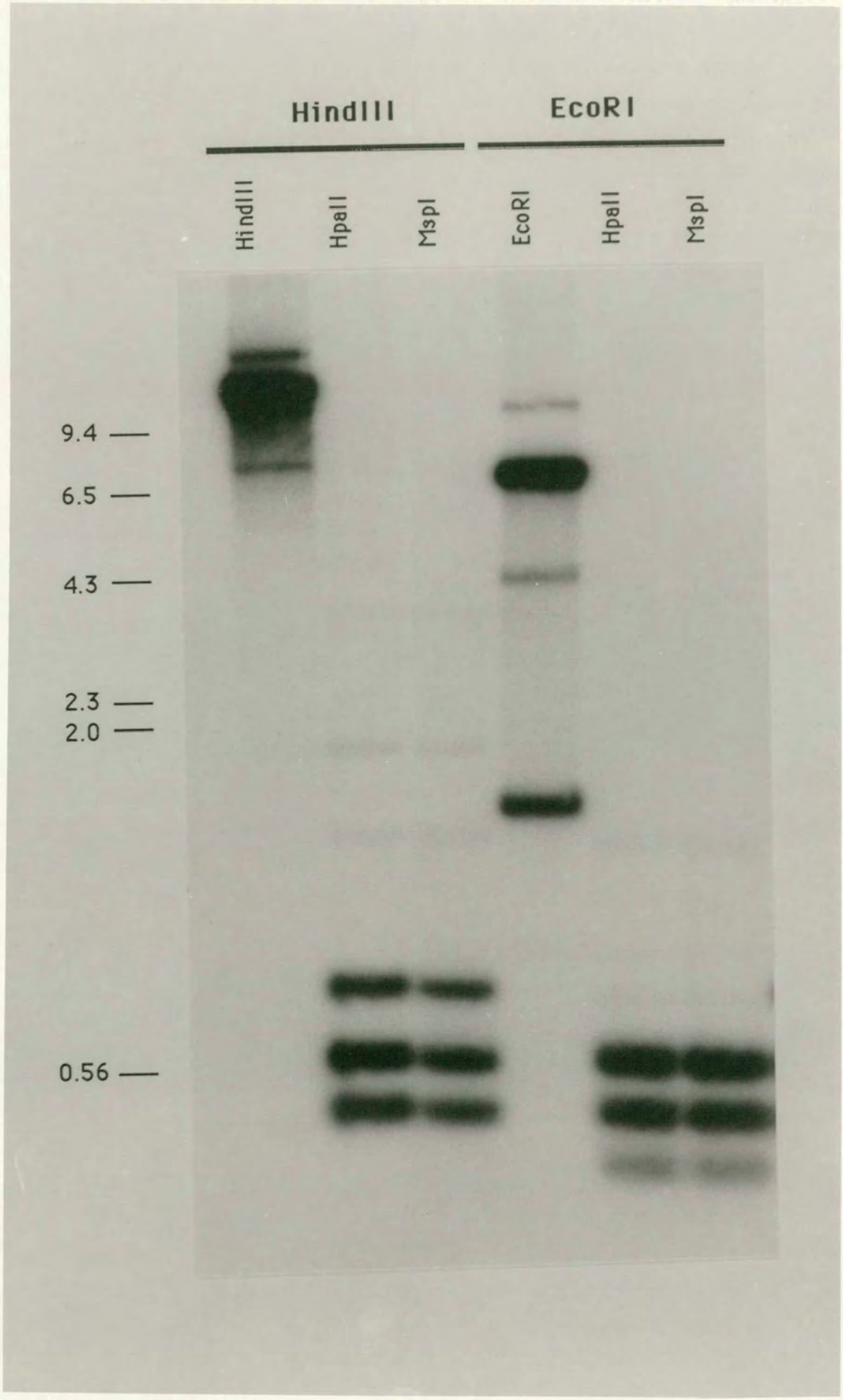

HindIII
EcoRI
HindIII
HpaII
MspI
EcoRI
HpaII
MspI
9.4
6.5
4.3
2.3
2.0
0.56

Figure 3.6

A Methylation analysis of repeated sequences in *S. pombe*

Southern blot of 1µg of *S. pombe* genomic DNA digested with restriction enzymes as indicated and hybridized to a probe made from 50ng of *S. pombe* genomic DNA. The molecular weight of the genomic DNA used as a probe was reduced by digestion with *MspI*. Film was exposed overnight at -70ºC. Molecular weight standards are indicated in kilobases.

B Methylation analysis of the *S. pombe* transposable elements

Southern blot of 1µg of *S. pombe* genomic DNA hybridized with a probe made from the long terminal repeat of the *S. pombe* transposable element. Film was exposed for 2 days at -70ºC. Molecular weight standards are indicated in kilobases.

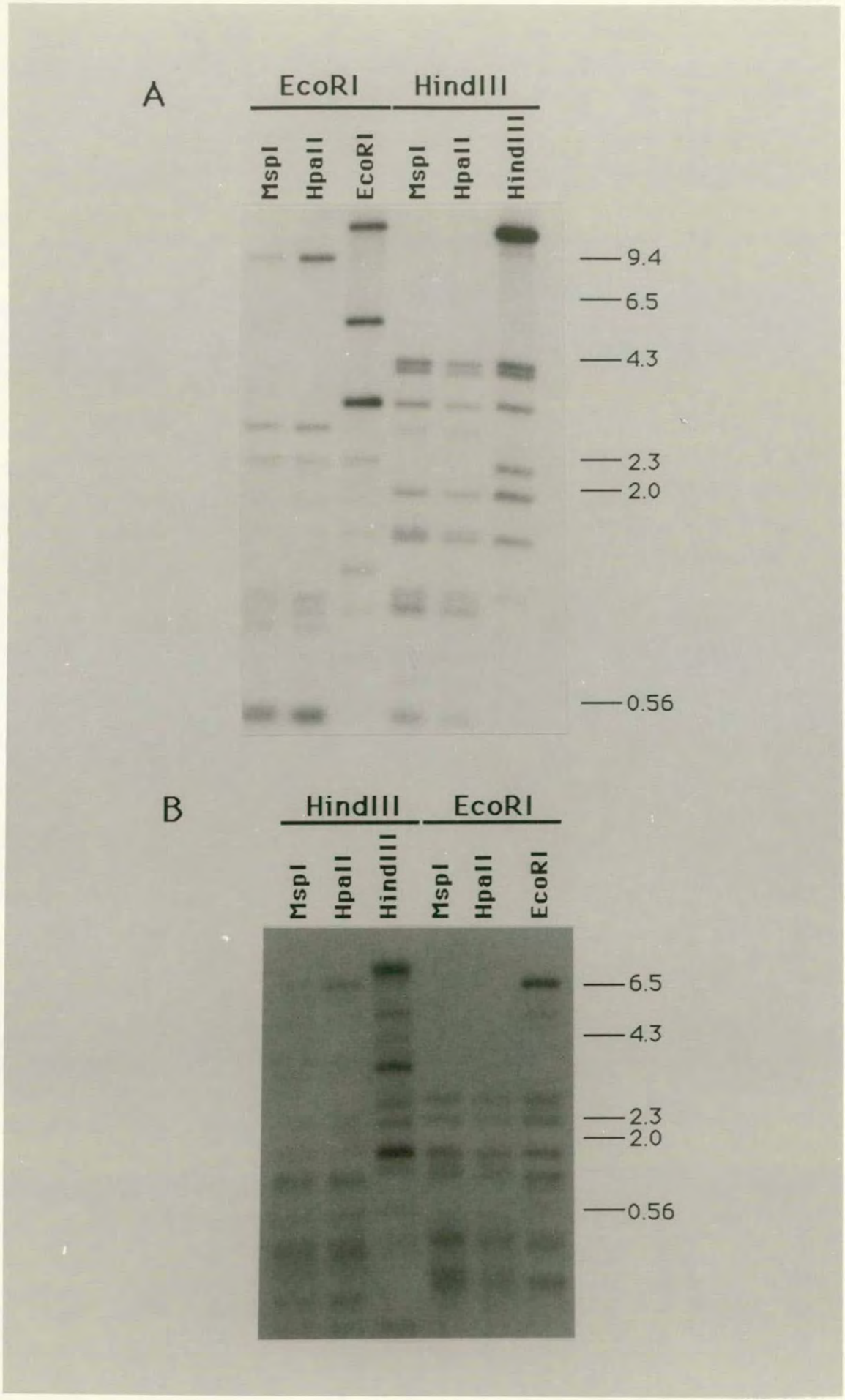

A
EcoRI
HindIII
MspI
HpaII
EcoRI
MspI
HpaII
HindIII
9.4
6.5
4.3
2.3
2.0
0.56
B
HindIII
EcoRI
MspI
HpaII
HindIII
MspI
HpaII
EcoRI
6.5
4.3
2.3
2.0
0.56

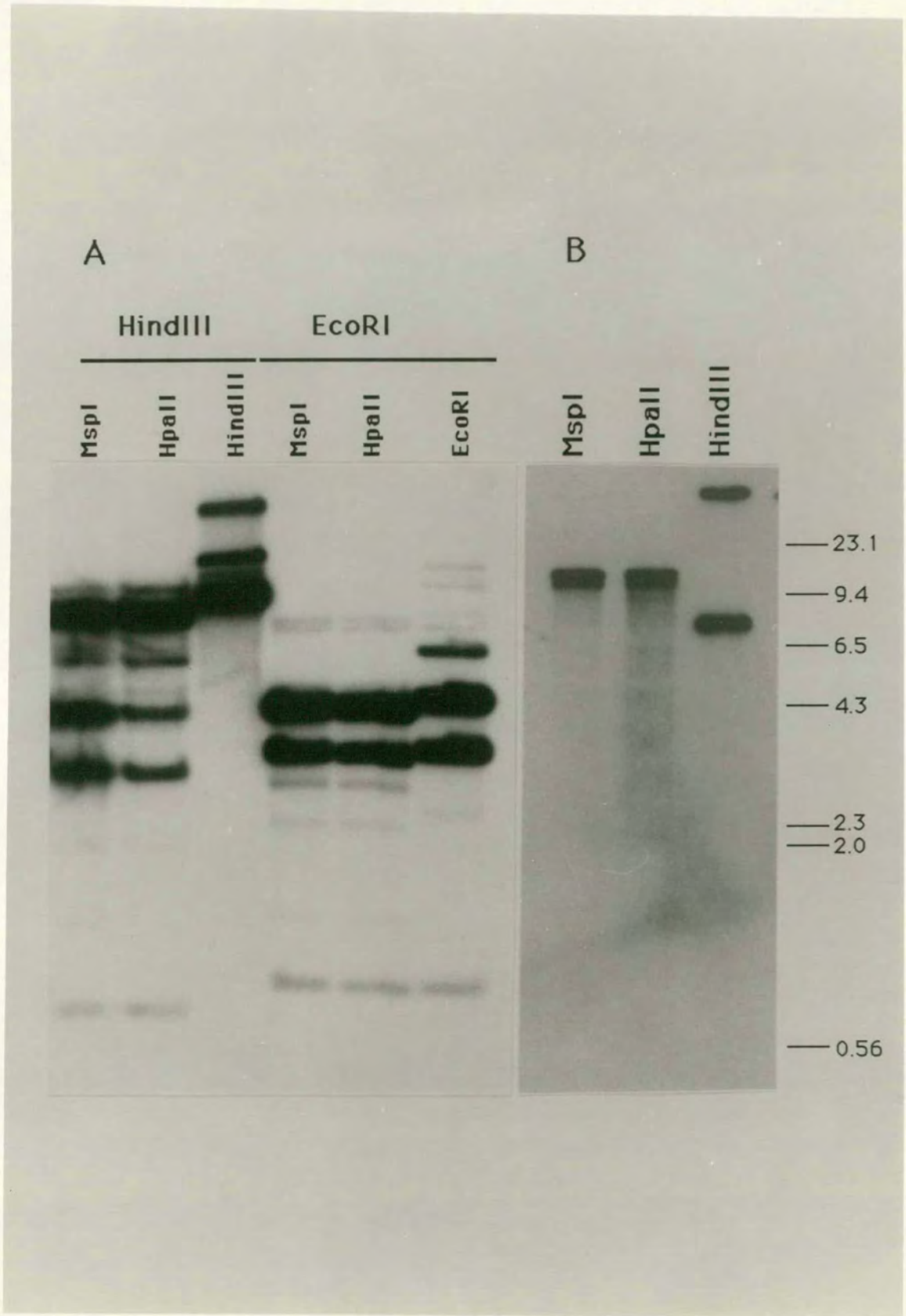
Figure 3.7

A Methylation analysis of the dgdh centromeric repeats in *S. pombe*

Southern blot of 1μg of *S. pombe* genomic DNA digested with restriction enzymes as indicated and hybridized with a probe made from the 6.2kb *EcoRI* fragment containing the *S. pombe* dg and dh centromeric repeats. Film was exposed overnight at -70ºC. Molecular weight standards are indicated in kilobases.

B Methylation analysis of the telomeric regions in *S. pombe*

Southern blot of 1μg of *S. pombe* genomic DNA hybridized with a probe made from the 7.1kb *HindIII-EcoRI* fragment containing the *S. pombe* telomeric-repeat sequences and telomere-associated sequences. Film was exposed at -70ºC overnight, molecular weight standards are indicated in kilobases.


A
HindIII
EcoRI
MspI
HpaII
HindIII
MspI
HpaII
EcoRI
B
MspI
HpaII
HindIII
23.1
9.4
6.5
4.3
2.3
2.0
0.56

identify repeated regions of *S. pombe* DNA containing m^5C. Cross *et al.*, (1994) have reported the construction of a column that is able to fractionate DNA according to its degree of methylation at CpG dinucleotides. The column contains the methyl-CpG binding domain from the rat MeCP-2 protein attached to a nickel agarose column by a histidine tag. DNA is loaded onto the column and then eluted with a linear gradient of NaCl concentration of between 0.25 and 0.8M. Non-methylated DNA is washed off at low salt concentration (0.4M NaCl) whereas methylated DNA is washed off between 0.65 and 0.8M NaCl according to the extent of methylation. The only sequence specificity requirement of this column is that m^5C is part of a CpG dinucleotide. Each MBD molecule is capable of interacting with a single CpG pair although increasing levels of methylation will result in tighter binding to the column presumably due to interactions with multiple MBD moieties.

Fission yeast DNA was applied to the column and the resulting fractions analyzed by end-labelling and agarose gel electrophoresis (Figure 3.8). No bands were identified in the fraction corresponding to higher salt concentrations where methylated DNA would be expected to appear. This indicated that there is not a significant methylated DNA fraction in *S. pombe* DNA consistent with previous reports (Antequera *et al.*, 1984).

3.5 Conclusions

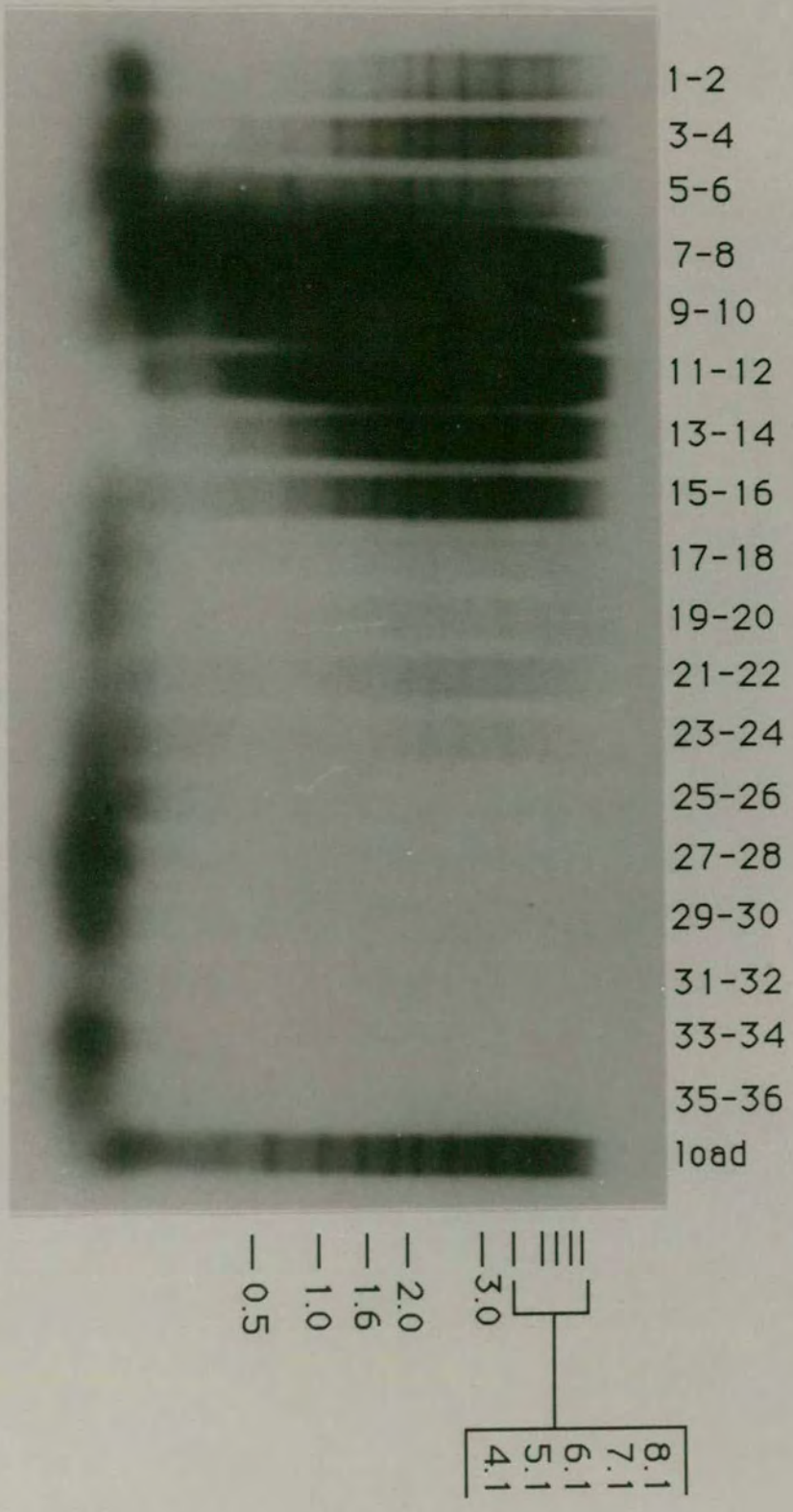
This chapter describes the search for CpG methylation in the *S. pombe* genome using the isoschizomeric restriction endonuclease analysis of specific regions. *S. pombe* DNA was also applied to a methylated DNA binding column in an attempt to detect a methylated fraction. In addition, an assessment of CNG methylation using nearest neighbour analysis of *BstNI-*

digested DNA was carried out. Despite the use of a range of techniques, it has still not been possible to detect m^5C in the *S. pombe* genome.

This analysis was carried out in attempt to identify regions of DNA that might be methylated by the cnd1 protein. The nature of the search was influenced by the distribution of DNA methylation in other organisms and by the phenotypes associated with the *cnd1-1* mutation. Since the experiments described were unable to shed any light on the possible functions of cnd1, different lines of investigation were chosen and the resulting studies will be described in the following chapters.

Figure 3.8 Fractionation of *S. pombe* DNA by a methylated DNA binding column

S. pombe DNA (10μg) that had been digested with *EcoRI*, was passed over the methylated DNA binding column. The column was washed with a linear gradient of NaCl (from 0.4M to 0.8M). A total of 36 fractions were collected consisting of 2mls each. The fractions were pooled pairwise (200μl of each), ethanol precipitated and then end-labeled in the presence of [α-^{32}P] dATP. Labelled DNA was separated on a 0.8% agarose gel which was then dried down and exposed to film overnight at -70°C. The numbers of the pooled fractions are indicated across the top of the gel and the sizes of molecular weight standards (in kilobases) are also shown. The fractions 25-36 correspond to the NaCl gradient between 0.65 and 0.8M and would be expected to contain any DNA that contained m^5C in CpG dinucleotides.


1-2
3-4
5-6
7-8
9-10
11-12
13-14
15-16
17-18
19-20
21-22
23-24
25-26
27-28
29-30
31-32
33-34
35-36
load
0.5
1.0
1.6
2.0
3.0
8.1
7.1
6.1
5.1
4.1

Chapter 4: Sequence analysis of *scd1+*

4.1 Introduction

Why sequence the suppressor of a mutation? The sequence analysis of a multicopy or extragenic suppressor can often provide information or clues relating to the function of the original gene. For example, the suppressing gene could encode a substrate of the gene product in question or a protein that is required for the activation or repression of the relevant gene activity. Alternatively, the identity of a suppressor may help to place a particular gene function on a certain pathway or into a particular cycle of events. Such information may be revealed by the homolgy of the suppressor's gene product to other sequences in the database or by the existence of certain diagnostic motifs within a protein sequence.

A multicopy suppressor of the *cnd1-1* mutation was isolated at the same time as the *cnd1* gene was cloned (Bartlett, 1991). I decided to sequence this suppressor in an attempt to gain an insight into the function of the *cnd1* gene. The sequence analysis is described in this chapter.

4.2 The suppressor of the *cnd1-1* mutation

The plasmid psup1-1 was isolated from a *S. pombe* genomic library as a multicopy suppressor of the conditional lethal phenotype of the *cnd1-1* mutation. The rescuing activity encoded by this plasmid was named *scd1+* for suppressor of *cnd1-1* (Bartlett, 1991). Mutant cells that are transformed with psup1-1 are able to grow at 36ºC and appear to be wild type by microscopic examination.

The library from which this plasmid was isolated was made by a partial *SauIIIA* digest of genomic DNA inserted into the *BamHI* site of the vector pYep13. The rescuing plasmid was found to contain a 6.8kb insert. Preliminary restriction enzyme mapping allowed the construction of psup1-2 which consists of a smaller 3kb *EcoRI-SalI* fragment of the psup1-1 insert subcloned into the vector pUR18. Psup1-2 is also able to rescue *cnd1-1* cells and thus the 3kb *EcoRI-SalI* fragment was predicted to contain the functional region of the *scd1*$^+$ gene (Bartlett,1991).

4.3 Sequence analysis of the *scd1*$^+$ genomic clone.

i Experimental strategy

The original restriction map of the psup1-1 plasmid (Bartlett, 1991) was found to be incorrect in that the indicated *EcoRI* sites were in fact *SalI* sites and *vice versa*. A corrected version of this map is shown in Figure 4.1A. The 3kb *EcoRI-SalI* insert of psup1-2 was subcloned into pBluescript KS$^-$ (Stratagene) to give the plasmid pBsup1-2. Sites for a number of restriction enzymes within the 3kb insert of pBsup1-2 were mapped (Figure 4.1B).

To prepare a series of subclones suitable for sequencing, nested deletions were made of pBsup1-2 following the method of Henikoff (1984), as described in Materials and Methods (section 2.7 i). The rate of deletion by exonuclease III was adjusted to give deletions of approximately 300 nucleotides spanning from the *SalI* site and the *EcoRI* site (Figure 4.2). These subclones were sequenced using the dideoxy chain termination method (Sanger *et al.*, 1977; section 2.7. ii) and the resulting sequence information assembled using the MacVector sequence analysis software. Gaps in the composite sequence were filled using sequence generated from reactions

Figure 4.1 A Restriction map of psup1-1.

The positions of the *EcoRI* (E) and *SalI* (S) sites are indicated as is the 3kb region that was subcloned to form psup1-2. Arrowheads show the *BamHI* site into which the *SauIIIA*-digested insert DNA was ligated in the original library construction.

B Restriction map of the 3kb insert of psup1-2.

Enzymes that were found not to cut include *BamHI, XhoI, KpnI* and *SmaI.*

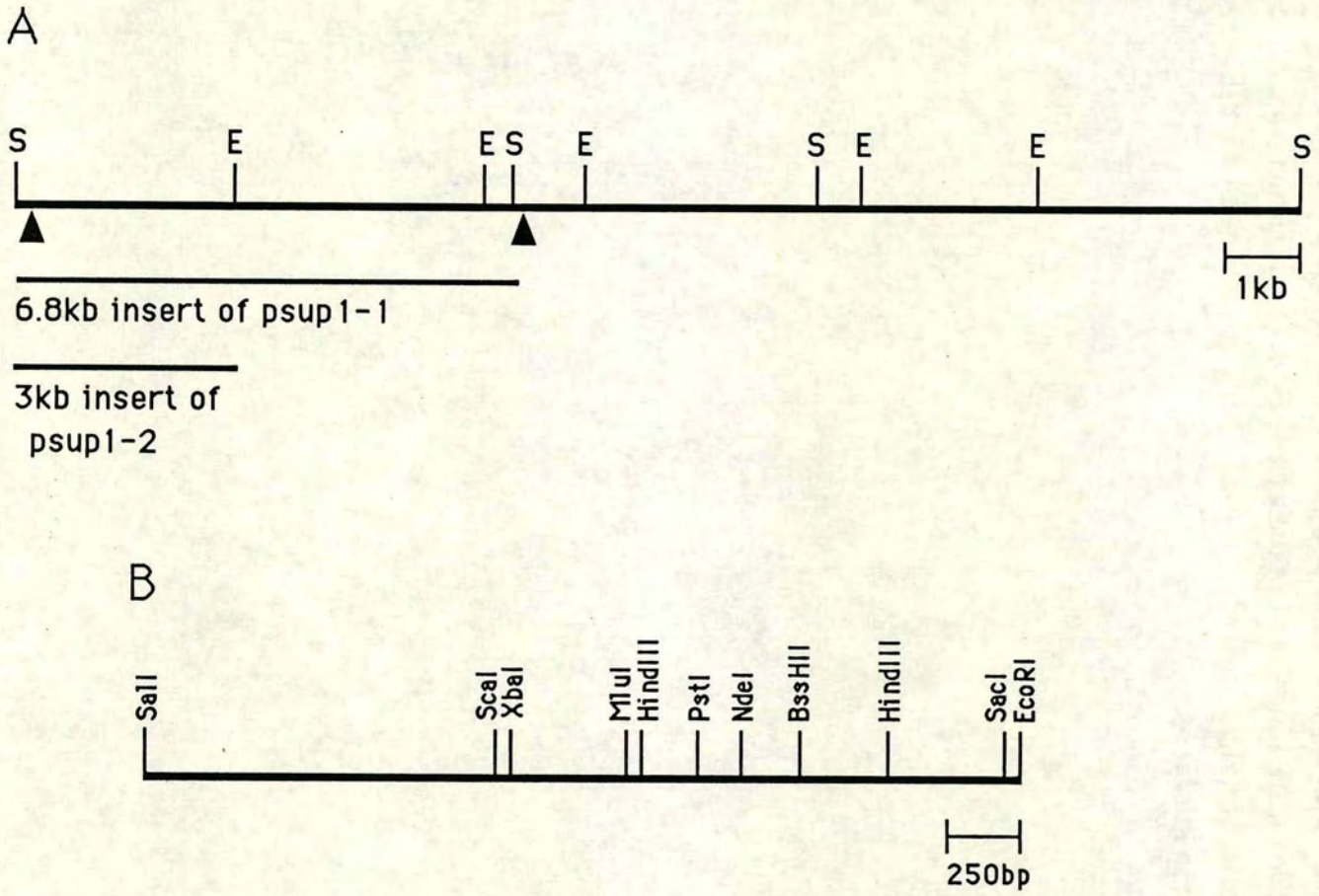
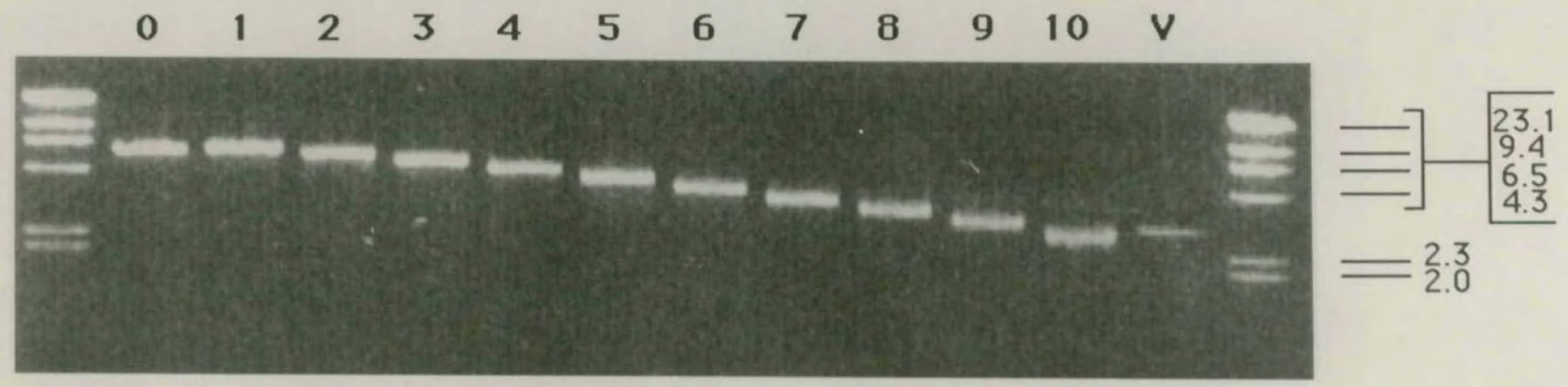

A
S
E
E S
E
S E
E
S
1kb
6.8kb insert of psup1-1
3kb insert of
psup1-2
B
SalI
ScaI
XbaI
MluI
HindIII
PstI
NdeI
BssHII
HindIII
SacI
EcoRI
250bp

Figure 4.2 Exonuclease III deletion series of pBsup1-2.

Exonuclease III deletions of pBsup1-2 of approximately 325 nucleotide incremental size (lanes 1-10), analyzed on a 0.8% agarose gel. The lane numbers correspond to time of incubation (in minutes), with exonuclease III. Linearized pBluescript KS^- vector is shown as indicated (V). Molecular weight standards (lambda DNA digested with *HindIII*) are shown in kilobases.


0
1
2
3
4
5
6
7
8
9
10
V
23.1
9.4
6.5
4.3
2.3
2.0

primed by synthetic oligonucleotides (*scd1*A-D, see Appendix A). The positions and orientations of deletions, oligonucleotides and subclones used in the sequencing of pBsup1-2 are summarized in Figure 4.3.

The insert of pBsup1-2 was found to be 2944 base pairs long (Figure 4.4). Confirmation that the sequence had been assembled correctly was provided by the presence and distribution of restriction enzyme sites that had previously been mapped. The 275 nucleotides immediately downstream of the *SalI* site are derived from the vector pYEP13 and this sequence is not shown in Figure 4.4. This sequence was incorporated into psup1-2 and subsequently into pBsup1-2 as the *SalI* site used for subcloning lies 5' to the *BamHI* site into which the *SauIIIA*-digested yeast DNA was ligated (Figure 4.1). A probe made from the 1.7kb *XbaI-EcoRI* fragment of pBsup1-2 was hybridized to a Southern blot of *S. pombe* genomic DNA cut with *EcoRI* and *XbaI*. A single band was seen at 1.7kb indicating that the insert of pBsup1-2 is not grossly rearranged between these two sites (data not shown).

ii Analysis of the open reading frame

Computer analysis of the pBsup1-2 sequence revealed that there was one major open reading frame (ORF) of 705 nucleotides starting 80 nucleotides upstream of the *MluI* site. The putative start codon (ATG) most likely to initiate translation is indicated (Figure 4.4). The nearest upstream ATG is 930 nucleotides 5' to this sequence. The region between this upstream ATG (-930) and the putative start ATG (+1) contains numerous stop codons in all three frames. Thus it is unlikely that the major ORF is spliced to any other sequences upstream. There are two further ATGs within the open reading frame, located 165 and 237 nucleotides downstream from

Figure 4.3 Strategy for sequencing of the pBsup1-2 insert.

A Map of the pBsup1-2 insert. Arrows indicate the approximate length of each sequenced subclone. A-D indicate oligonucleotides that were used to prime sequencing reactions (see Appendix A for details). Numbers indicate subclones referred to in the text (1) *HindIII-HindIII*, (2) *BssHII-XbaI*, (3) *PstI-SalI*, (4) *HindIII-SalI*. All other arrows represent subclones generated by exonuclease III deletions. The region sequenced on both strands and the regions sequenced on one strand are represented by a thick and a thin line respectively.

B Restriction map of pBsup1-2 showing the predicted *scd1*$^+$ open reading frame represented by a thick arrow.

A

A
B
C
D
1
2
3
4
SalI
ScaI
XbaI
MluI
HindIII
PstI
NdeI
BssHII
HindIII
EcoRI
150bp

Figure 4.4 Nucleotide sequence of pBsup1-2 and predicted peptide sequence of *scd1*$^{+}$

Nucleotide sequence of pBsup1-2 sequenced on both strands between nucleotides -175 and 973 and on one strand in the rest of the sequence. Sites for the restriction enzymes shown in Figure 4.1B are underlined as are putative TATA boxes upstream from the *scd1* ORF. Predicted start and stop codons of *scd1* are shown in bold typeface. The nucleotides are numbered relative to the predicted start codon and the putative start and stop codons are shown in bold typeface.

```
-1315                          GATCATTAAAGGCGTATCAACAGTGTTGGGGAAA  -1283

-1282  AAGTTTACCTTGAATCTTATCTAGCAGCGGAAACTAAATAATTATATTTAT        -1232

-1231  TAAATTGTTAGACCAGCCTAAAACTATCGGGATTAGCGTTGGGTAAGGGAT        -1181

-1180  TTACGGAATATTTACCGTATTAGGTTTTTTTTGAGTTTTAAAAATTTAACT        -1130

-1129  CGCTAACATAACCGAAGCCTCGCACATATAAACTATCTTTCAACCTACTTG        -1079

-1078  ACTCAACTTTTCCTTCACCAACCGCAGTGTTGCATTAAGCAGCAAAACGTT        -1028

-1027  GTCACCTTCGAATAGCGTCTTTTTGACGAATAAAAAGGTTGGACAACAAAA         -977

 -976  TTAAATTAAGCAAATTTGTGTATAAACTAACAGCAAGTGAAGTTGTACCGG         -926

 -925  TTTGTATGATTGAAGCATTGTAAACCAAAGAACAGCTACTTTGCCTTGTCG         -875

 -874  TACATTAGCTTTTTCCTTTGCAAAGAACTTTTAGGGACCTTTTTTTTACCAC        -824

 -823  AACTAGAGAACCGGTGTAGAAAGGTGTTCTATTTTAGTGAACGGATTCGTT         -773

 -772  TAATTTTTTCATTGAGAAAAAAATTTGATCTATCTTGATTTTTTCGACTGT         -722

 -721  TTCCATTTTTAACTATTTGTGACGTTTCTTATTGCCAAATTTCAACCTTGA         -671

 -670  GCATTGTTAAAAAAGCTGAATTTAAATTTCTTTGCTTTTAACATACAAGTG         -620

 -619  AGCTATTGGGTTTGTAAAGATTTAATACTTGATTGGAAGTTTTTTTGGTCA         -569

 -568  AGTCTGTTTTTTTTGGGTGTTTTTTTGACGATCAGTTGAAACCGCGAATAA         -518

 -517  TAATATTTATTATAACTCCCTAAAAAAAGAAAGAAAATCATTTATTAAATA         -467

 -466  ATTTTTACTTTCCTTCATTGTTTCTGAATTTTGCTTCTCCCTGCCGTTTAA         -416
                                  ScaI
 -415  TTGTTTCATTTTCTTTTAGAAAGTACTAAAAGGTAATCAAATCGGCTAGTC         -365
```

XbaI

-364 TTGCAACTGAATTTTCTTTGAAGATAATATCTAGATTTAATTTTCCCTGCT -314

-313 TGAATTTCCTAAGATCTTGCGTTCGTTCGGACTGAATTGCATAATTTATTG -263

-262 AAGGTCTCCATTCCTGTATACTTACTCGACTACCCATTTATCGGTCAAGCT -212

-211 GTATATATACACGCACAGTCAACTTATCTTCTACTTATATTTATAAAACTA -161

-160 TTTATTTTATTTCTCTTTAATCCGAGCTTCGTTAGTTGATTCTTTTGATCT -110

-109 TTGCATTGGTAACTAATCCAGTGCGCTTTTTTTTTTGAATCTTCGTGACAG -59

-58 CTATTTTCTGATTTTACTTCATCTTCAAACTACGATCTGATCGGCTTTGCC -8

-7 TTTCGTC**ATG**CAAACTTATACAGGTATCATTAAGACCCCGTTGGATGCGAT 4
M Q T Y T G I I K T P L D A I

MluI

44 TATCCTTTTCGAGGCCTGTCGAATTGGTCTTCTTCCACGCGTCCAGCGTCG 94
I L F E A C R I G L L P R V Q R R

HindIII

95 ATTGTCCGACCATGAAAGAAGCTTGATACGAGCAGGCAGTGTTTTTGTGTG 145
L S D H E R S L I R A G S V F V W

146 GGATGAACGGGAAGCTGGAATGCGCCGTTGGACTGATGGTAAATCATGGAG 196
D E R E A G M R R W T D G K S W S

197 TGCTTCTCGTGTCAGCGGTTCATTCTTAACCTACCGAGAAATGGAGGGTAA 247
A S R V S G S F L T Y R E M E G K

248 ACGTAAACCTTACCATCACGGTCTTTCCACCGATGGTTCCCTTAAACGCTC 298
R K P Y H H G L S T D G S L K R S

PstI

299 TCCCTCTGCAGATACGACTGGTAATAGTTCCCTAAATGCTTACAGCAATGA 349
P S A D T T G N S S L N A Y S N E

350 GGACAGTGGAGCCGCGAGCTTGAGTGATGAAGAGTCGGTTGATGATGAGAA 400
D S G A A S L S D E E S V D D E N

401 CCTACGGGGGTTGCATTACAAACCGAACGGCTTAATTAAGCAGTCCTTTAG 451
L R G L H Y K P N G L I K Q S F S

452 TATTACTACCTCCTTAAATCATAAACTTCACCTTATTTCCTATTCTTCGCC 502
I T T S L N H K L H L I S Y S S P

503 TATCCCTGATCCTTCATTAGTCACGCCAAGTAGTGATGTTAACCTTTCACG 553
I P D P S L V T P S S D V N L S R

NdeI
554 TATCACTATTCCATATGGACTGTATCCAGACGCGGGTCCACCGCTTGTTCA 604
I T I P Y G L Y P D A G P P L V Q

605 AGCACCAAAGTTTCTCGCTCCTTATGATATTCTTGTCGAGAAAGTTGCCTG 655
A P K F L A P Y D I L V E K V A C

BssHII
656 TAGCGAAGACGCGCGCGTTTTGGATAAGTTGCGTCAAGCCCTATGGCTATA 706
S E D A R V L D K L R Q A L W L

707 GGTTTGATCTTTTTCTTCACATTGGGTCCTACTTCTCGGTTAAAAAGTGTA 757

758 CTAAATTTACTATAGATTTCATTTGCATCATATATGGGGCTTTAACGACTC 808

809 CCAATCAATCCTAGTTTCCTGTGACATTTTCTCGGCTGCGTTTTCGCTCTT 859

860 ACTAATCTGCGGCTTTGTCTATCAGCGTCTAATTTATGTGGTTGATCAATT 910

911 TTAATTTCAAGTTACCACTTCATTCTTCAGCACTTACGGCAAGTGATAAAA 961

HindIII
962 ACAGAAAGCTTGTTTTGATTTAATTCATTAACTAAGTAATCGCTCAACTGT 1012

1013 CTTGCAAATACGTTTTCTGTTCGTCGGTTTAGATGTGTATCGTTATCATAA 1063

1064 CTCTTTAGACATTTCGCTTTACGCATTATTTTGTTTCGCTGCGAATATCGG 1114

1115 AAATGTCCTTTTTAAAGAATCTTTTGTGGTTATTTCATCAATGCTTCCTTT 1165

1166 TTTCTCTTTATTATTTATTTCTTTTCATGATGGAAATGGTTAAGTAAACGT 1216

1217 CTATCATGAAGCACGGTAGAACAAGCTATGCCAACATTTGAATTATGCACT 1267

SacI
1268 CAACATTTCTATTTTTTACTACTCAACTTAGAGCTCATGTTCGGGTAAAAT 1318

EcoRI
1319 GATAGGAAAGCATCAATTCTTAGGGTAATTTGAATTC 1354

the putative start ATG. However, translation has been found in the majority of cases, to start at the most 5' ATG codon of an ORF (Kozak, 1983).

The ORF terminates with the stop codon TAG as indicated (Figures 4.4). This TAG is likely to be the end of the putative *scd1*$^+$ gene as the 649 nucleotides downstream contain numerous stop codons in all three frames. The region encompassing the large open reading frame was sequenced on the opposite strand using subclones generated by restriction digestion as follows; 1) *HindIII-HindIII*, 2) *BssHII-XbaI*, 3) *PstI-SalI* and 4) *HindIII-SalI*. These subclones were sequenced on an ABI machine (section 2.7 iii) with the assistance of N. Tountas.

iii 5' sequence

Analysis of the 5' non coding region of the predicted *scd1*$^+$ gene reveals that there are several potential TATA boxes upstream of the predicted start codon (Figure 4.4). The TATA motif located 170 nucleotides upstream of this ATG is a perfect match to the higher eukaryotic TATA box sequence TATAAA. The locations of all the putative TATA boxes is in good agreement with that expected for *S. pombe* genes (within 250 nucleotides upstream of the ORF; Russell, 1989). In general, the consensus for *S. pombe* TATA motifs is not very strong, but TATA sequences can usually be found 35-45 nucleotides upstream of the transcriptional start sites of genes which are transcribed at moderate or high levels (Russell, 1989). No other consensus motifs for upstream regulatory elements were found in the region 5' to the putative *scd1*$^+$ gene.

iv 3' sequence

Most eukaryotic mRNAs of higher eukaryotes have a polyadenylated 3' terminus that starts approximately 20 nucleotides downstream of a transcriptional terminator sequence related to 5'-AAUAAA. The 3' ends of the mRNA is formed by endonucleolytic cleavage and polyadenylation of the nascent mRNA transcript by nuclear factors that interact with the terminator sequence. Many genes in budding and fission yeast lack the consensus terminator sequence altogether (Russell, 1989). In *S. cerevisiae,* an alternative terminator sequence 5'-TAT....TAGT....TTT-3' has been discovered that can be found up to 160 base pairs downstream of the stop codons (Zaret and Sherman, 1982).

The polyadenylation signals of *S. pombe* genes, like those of *S. cerevisiae* are AT-rich but possess no obvious sequence motif (Humphrey *et al.*, 1991). Analysis of the 3' non-coding region of the putative *scd1*$^+$ gene is consistent with these observations as there are no sequences corresponding to either the 5'-AATAAA-3' or 5'-TAG....TAGT....TTT terminator sequences, but the AT content of the DNA 100bp downstream of the TAG does increase by 12% compared with the preceding coding sequence.

v Analysis of the protein sequence

Translation of the putative ORF results in a predicted polypeptide of 235 amino acids (Figure 4.4). The homology search algorithm FASTA (Lipman and Pearson, 1985) was used to search the SwissProt database for proteins showing similarity to the predicted *scd1*$^+$ product. No homology was found to any of the entries in this database, nor were any known motifs identified using the PROSITE motif programme.

Analysis of the amino acid content of the predicted scd1 protein reveals that there are equal proportions of acidic and basic residues and that 22% of the amino acids are either serine, threonine or tyrosine, all of which are known to be commonly phosphorylated. Serine is present in the greater amount accounting for 13% of the amino acid residues. This finding suggests that scd1 may be a phospho-protein.

4.3 Conclusions

This chapter describes the sequence analysis of the insert from the plasmid psup1-2 which was predicted to contain the *scd1*$^+$ gene, a multicopy suppressor of the *cnd1-1* mutation. The *EcoRI-SalI* insert from psup1-2 was found to be 2944 nucleotides long and contained one large open reading frame which encodes a predicted protein of 235 amino acids long. Analysis of the DNA sequences 5' and 3' to this major ORF suggests that the *scd1*$^+$ gene consists of a single exon. The predicted scd1 protein does not share homologies to any of the entries in the current SwissProtein database or possess any motifs that would suggest a possible function for this gene.

As the sequence of the *scd1*$^+$ gene alone did not provide any clues as to the function of the DNA methyltransferase homologue cnd1, further analysis of *scd1*$^+$ was not carried out at this time. Instead, attention was turned directly to the *cnd1* $^+$ gene and these studies will be described in the following chapters. Additional information about *scd1*$^+$ was discovered at a later date though and this will be discussed in chapter 6.

Chapter 5: Purification of the cnd1 protein and analysis of its methyltransferase activity

5.1 Introduction

Although the sequence of the cnd1 protein suggested that it was a member of the m^5C-MTase family, its ability to methylate DNA remained to be demonstrated. To address this issue, the cnd1 protein was purified and used in a series of *in vitro* methyltransferase assays. This analysis is the subject of chapter 5. The purified protein was also used to generate polyclonal antisera and attempts to detect the cnd1 protein in fission yeast cell extracts are also described in this section.

5.2 Construction of pet6Hcnd1

In order to purify the cnd1 protein, the *cnd1* cDNA was subcloned into the bacterial expression vector pet6H (section 2.2 v, Figure 2.4) to generate the plasmid pet6Hcnd1. The vector contains a sequence coding for six histidine residues upstream of the coding site so that the recombinant protein will possess a histidine tag at its N-terminus. This enables the protein to be purified by the use of nickel-ion affinity chromatography. Recombinant protein expression is induced by the addition of isopropylthio-ß-D-galactosidase (IPTG), which switches on the expression of the T7 RNA polymerase gene via the inducible lacUV5 promoter.

Oligonucleotides cnd15 and cnd13 were used to amplify the *cnd1* cDNA and incorporate an *NcoI* site at the 5' end and a *BamHI* site at the 3' end to enable the product to be subcloned into pet6H (details of these

oligonucleotides and the PCR reaction can be found in Appendix B). As errors can be incorporated into DNA during PCR, it was necessary to check that the *cnd1* sequence had not been altered. The cDNA from one of the pet6Hcnd1 clones was subcloned into pBluescript KS$^-$ (by creating a blunt end at the *NcoI* site and then subcloning the *BamHI*-blunt ended cDNA into pBluescript KS$^-$ that had been cut with *BamHI* and *SmaI*). The sequencing reactions were primed by the oligonucleotides cndA-E (Appendix B) and the universal M-13 primer. No mutations were found in the *cnd1* sequence and thus the pet6Hcnd1 plasmid was used for the subsequent expression of the protein.

5.3 Purification of the cnd1 protein

i Induction of cnd1 expression

The pet6Hcnd1 plamsid was transformed into the bacterial strain BL21(DE3)pLys (section 2.9 i). Upon induction with IPTG at 37ºC, cnd1 protein was synthesized to levels that enabled its detection in total cell extracts by staining with Coomassie Brilliant blue dye (Figure 5.1). This protein was not present in extracts from cells that contained the pet6H vector alone. The size of the induced band was in agreement with the predicted molecular weight of the histidine-tagged cnd1 protein which is 38kD. The levels of cnd1 protein increased steadily over the first four hours of induction and then did not increase significantly beyond this time. For this reason, all subsequent inductions were carried out for four hours. Varying the temperature of induction did not alter the overall levels of expression (Figure 5.2 and 5.4). Approximately 2.5mg of cnd1 protein were

produced in four hours by one litre of cells when expression was induced at 37°C.

ii Solubility of the recombinant cnd1 protein

The next stage in the characterization of cnd1 expression was to determine whether the expressed protein was soluble. After induction the cells were pelleted and resuspended in a non-denaturing buffer containing 50mM NaCl (buffer A, section 2.8 ii). Cell extracts were prepared and divided into two fractions by centrifugation. The cellular debris and insoluble material were collected in the pellet fraction while the soluble material remained in the supernatant. When the two fractions were analyzed by SDS-PAGE, the cnd1 protien was found in the pellet fraction indicating that it was insoluble in buffer A (Figure 5.2).

Induction of expression at lower temperatures of 30°C and 27°C also resulted in the production of insoluble cnd1 protein. The levels of cnd1 protein produced at these lower temperatures were similar to those synthesized at 37°C (data not shown). Therefore an attempt was made to purify cnd 1 from the pellet fraction. Sometimes recombinant proteins are insoluble and form aggregates known as inclusion bodies. Aggregates containing cnd1 were prepared and a substantial initial purification achieved by the solubilization of contaminating proteins in high concentrations of sodium chloride and then with 6M urea (Figure 5.3). The intention was to dialyze the inclusion bodies so that the cnd1 protein would eventually become solubilized in 5M urea. Buffers containing up to 5M urea can be applied to the nickel agarose column. This approach was not successful, however, as the cnd1 protein was not soluble in concentrations of urea that were less than 8M (data not shown).

Figure 5.1 Time course of induction of the cnd1 protein

Cells containing either the pet6H vector (V) or the pet6Hcnd1 (P) plasmids were grown at 37ºC to an optical density at 600nm (OD_{600}) of 0.75. At this point, 20µl samples were removed from each culture and IPTG added to a concentration of 0.4mM. Further 20µl samples were taken 30, 60 and 120 minutes after the addition of IPTG as indicated. Cellular protein from all the samples were separated by gel electrophoresis on a 10% SDS polyacrylamide gel. The gel was subsequently stained with Coomassie Brilliant Blue dye. In this and all subsequent figures in chapter 5, the arrowhead or arrow refers to the expected position of the cnd1 protein according to its predicted molecular weight (38kd). Molecular weight standards are indicated in kilodaltons (this also applies to all subsequent figures in chapter 5).

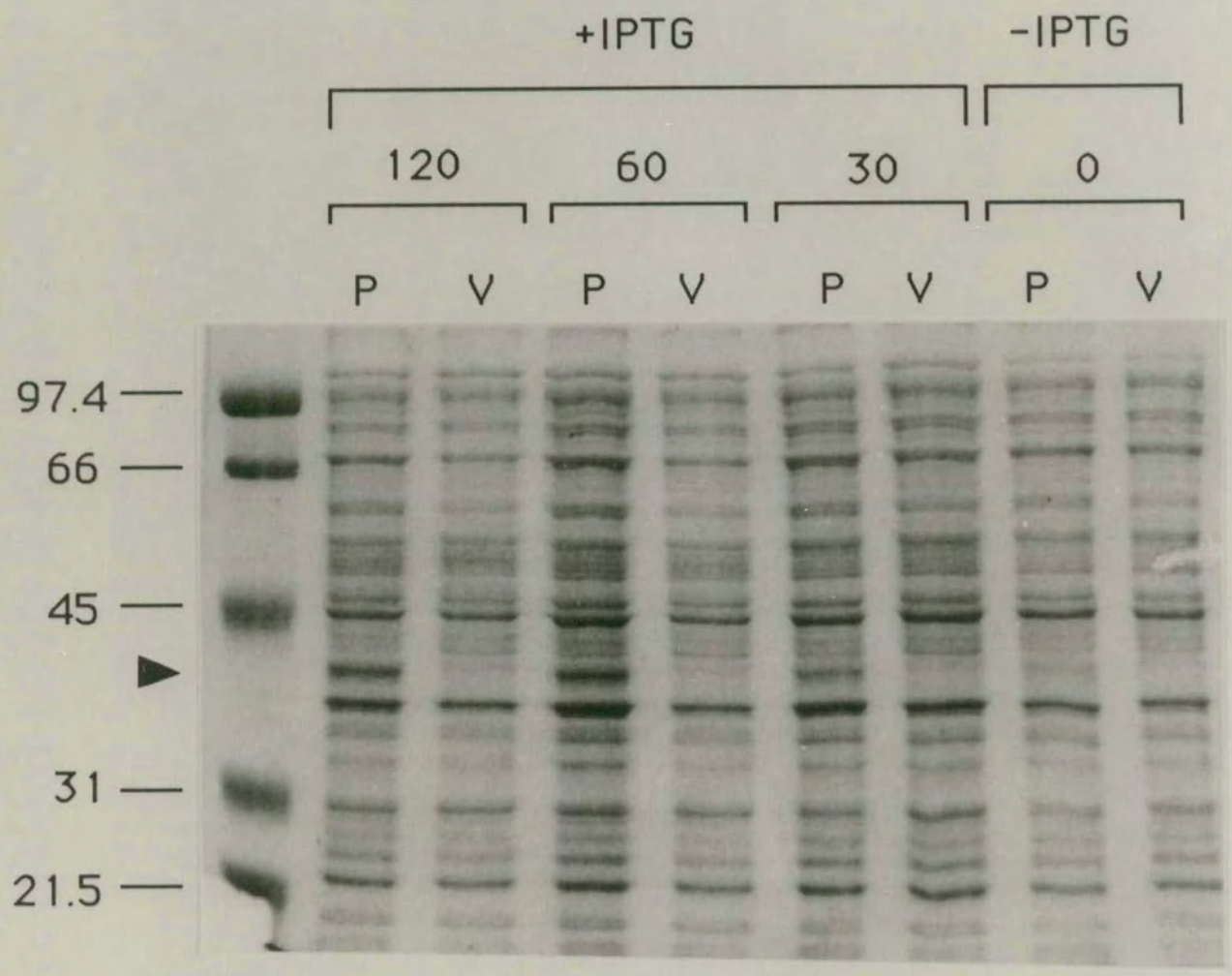

+IPTG
-IPTG
120
60
30
0
P
V
P
V
P
V
P
V
97.4
66
45
31
21.5

Figure 5.2 The cnd1 protein is insoluble when expression is induced at 37°C

Cells containing either the pet6H (V) or the pet6Hcnd1 (P) plasmids were grown at 37°C to an OD_{600} of 0.46 and then IPTG added to 0.4mM. After induction for 1 hour, cells were pelleted and resuspended in low salt buffer A (section 2.8 ii), and then insoluble cellular material removed by centrifugation at 16K for 30 minutes to give the pellet fraction. The supernatant (SUP) from this centrifugation step is also shown. Each lane contains extracts from 40µl of cells except the SUP-V lane which was overloaded. Samples were separated on a 10% SDS polyacrylamide gel which was subsequently stained with Coomassie blue.

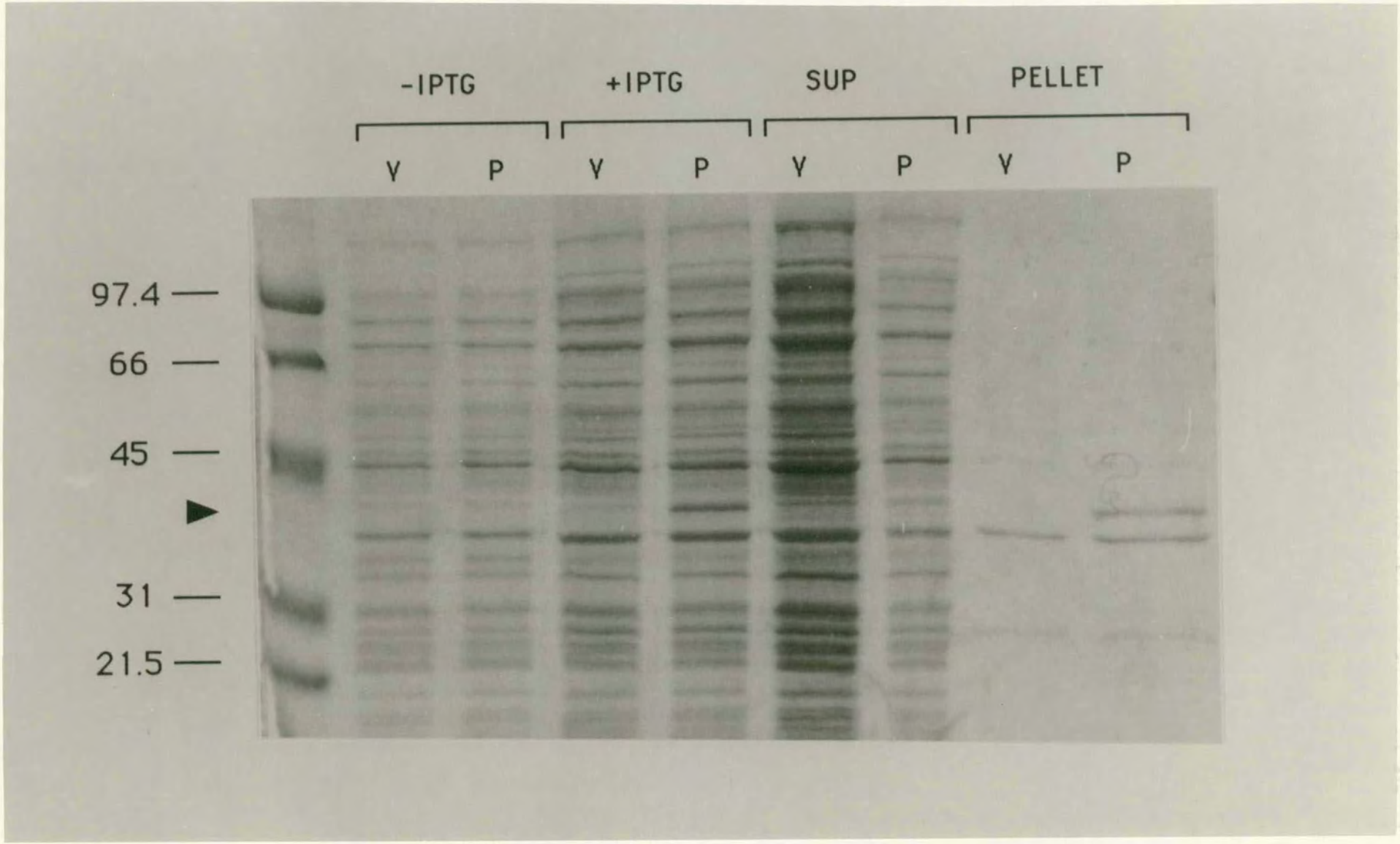
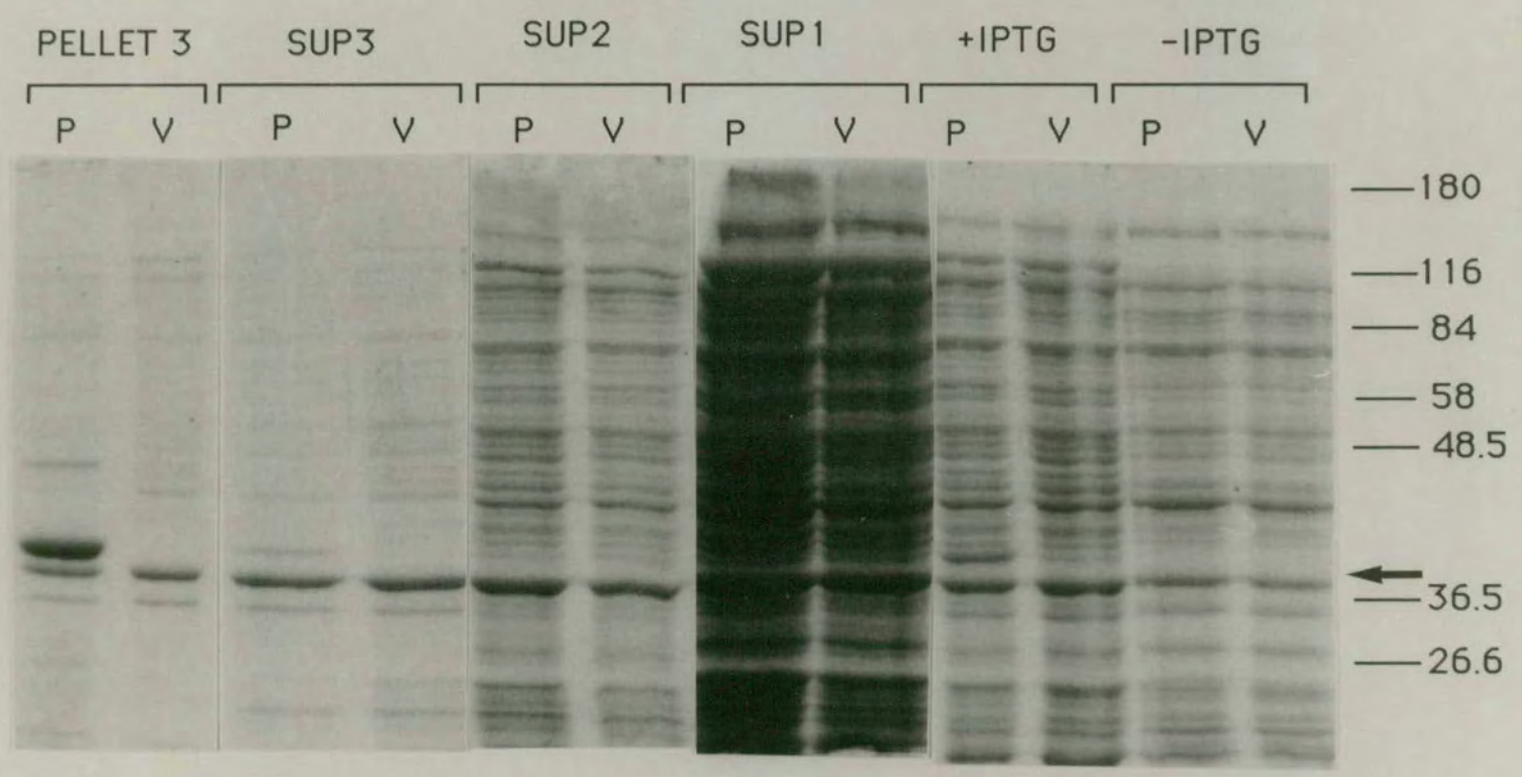

-IPTG
+IPTG
SUP
PELLET
Y
P
Y
P
Y
P
Y
P
97.4
66
45
31
21.5

Figure 5.3 Preparation of inclusion bodies containing the cnd1 protein

Cells containing either the pet6H (V) vector or the pet6Hcnd1 (P) plasmids were grown to an OD_{600} of 0.5 at 37°C and then induced with IPTG for 4 hours. Bacterial extract was prepared in buffer A containing 500mM NaCl as described in Figure 5.2. The insoluble cellular material was removed by centrifugation leaving the supernatant (SUP1). The cell pellet was resuspended in buffer A (500mM NaCl) and the suspension separated into supernatant (SUP2) and pellet. This pellet was resuspended in buffer B (6M urea, 500mM NaCl) and the resulting suspension separated by centrifugation to give the supernatant (SUP3) and the pellet-3. The pellet was resuspended in loading buffer (section 2.8 i). The lanes - /+ IPTG contain extracts from 60µl of cells. All supernatant and pellet lanes contain extracts from 200µl of cells. Samples were separated on a 10% SDS polyacrylamide gel.


PELLET 3
SUP3
SUP2
SUP1
+IPTG
-IPTG
P
V
P
V
P
V
P
V
P
V
P
V
180
116
84
58
48.5
36.5
26.6

It has been reported that low temperatures of induction can favour the solubility of recombinant proteins (Schein, 1989). Although temperatures of 30ºC and 27ºC had not proved successful, a further attempt to produce soluble cnd1 protein was made by inducing expression at 24ºC. Upon analysis of the resulting extracts, approximately half of the induced cnd1 protein was found to be soluble in 50mM NaCl (buffer A). The solubility did not apear to be affected by salt concentration (Figure 5.4) or length of induction (data not shown). Thus the supernatant from these extracts was used as the starting material for the purification step.

iii Nickel-ion affinity chromatography

Extracts containing the soluble cnd1 protein were passed over a nickel agarose column (Novagen). The column was washed extensively with buffer A and then histidine-tagged proteins were eluted from the resin with imidazole which is an analog of histidine. Buffer A containing 5 and 10 mM imidazole was used to remove weakly bound proteins. The cnd1 protein was eluted with buffer A containing 80 and 100 mM imidazole and the purified protein assayed using SDS-PAGE (Figure 5.5). A number of bacterial proteins that presumably contain a natural tract of histidine residues in their sequence, were found to co-elute with the cnd1 protein in the 80mM imidazole fractions and in the first two 100mM fractions. These proteins were also present in the same fractions prepared from cells expressing the pet6H vector alone indicating that the lower molecular weight bands were bacterial proteins rather than degradation products of cnd1 (Figure 5.6).

Figure 5.4 A proportion of the cnd1 protein is soluble when expression is induced at 24°C

Cells containing either the pet6H vector (V) or the pet6Hcnd1 plasmid (P) were grown at 24°C to an OD_{600} of 0.54 and then induced with IPTG for 4 hours. After induction the pet6Hcnd1 culture was divided into three fractions (Pa, Pb, Pc). These cells were resuspended in the following buffers: (a) buffer A containing 50mM NaCl, (b) buffer A containing 200mM NaCl, (c) buffer A containing 500mM NaCl. The cells containing the pet6H vector alone were resuspended in buffer A containing 50mM NaCl. The extracts were divided by centrifugation to give the pellet (P) and supernatant (S) fractions. Samples were separated on a 10% SDS gel. All lanes contain extracts from 20µl of cells.

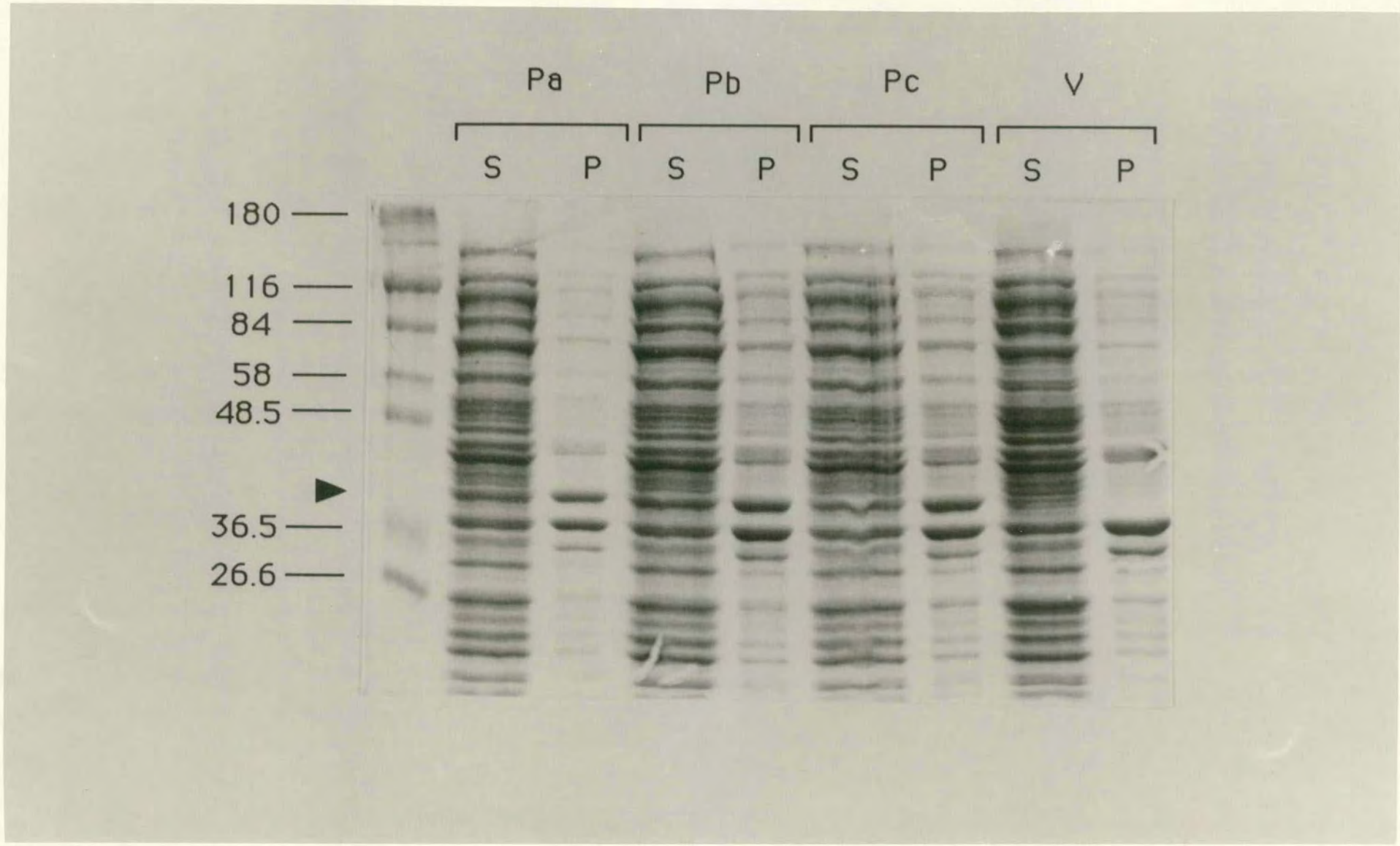
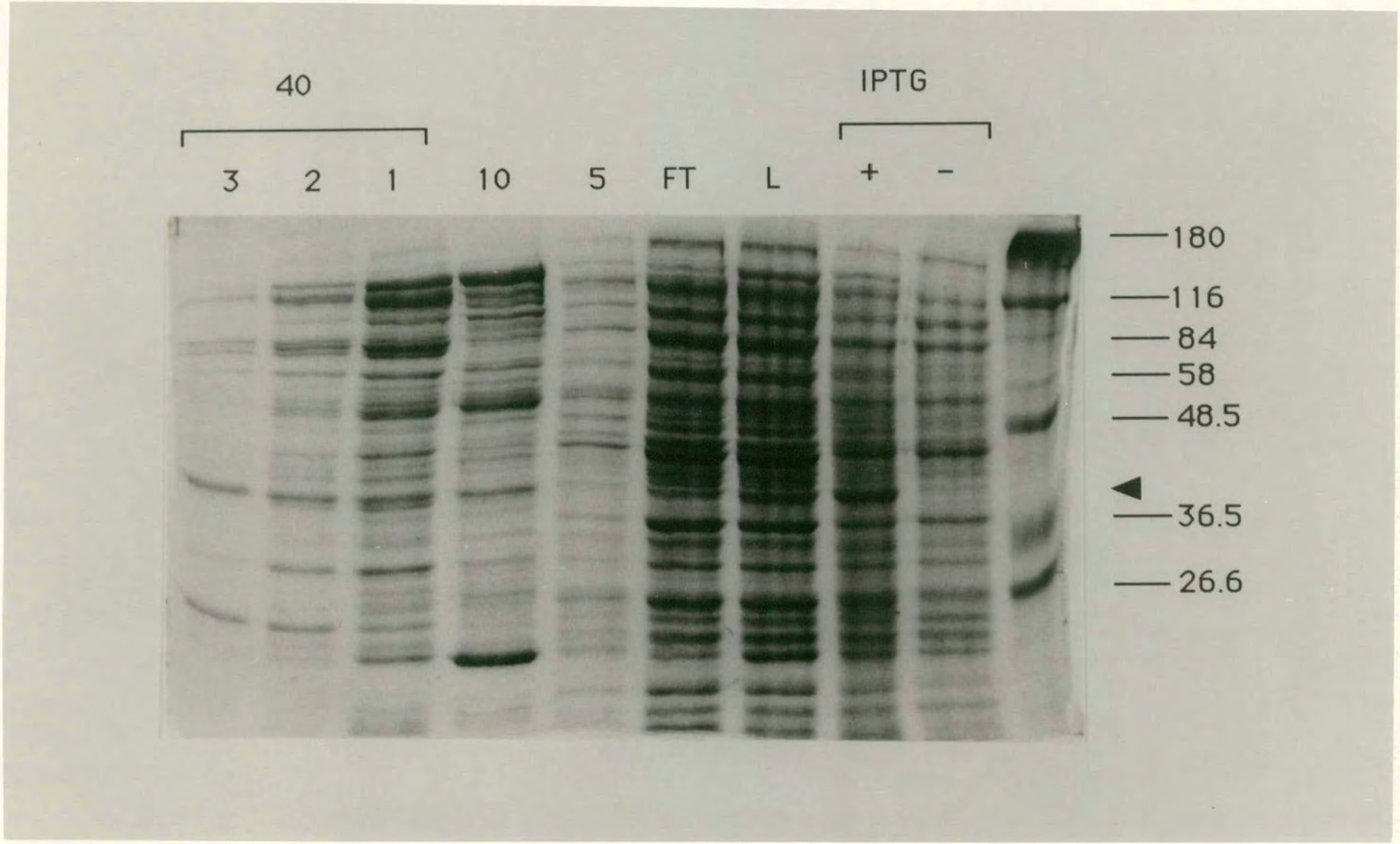

Pa
Pb
Pc
V
S
P
S
P
S
P
S
P
180
116
84
58
48.5
36.5
26.6

Figure 5.5 Purfication of the cnd1 protein by nickel-ion affinity chromatography

A Cells containing the pet6Hcnd1 plasmid were grown to an OD_{600} of 0.56 and IPTG added to 0.4mM for 4 hours. Cell extracts were made in buffer A and passed over a nickel agarose column as described (section 2.8 ii). L=column load; FT=column flow through. 5, 10 and 40 = buffer A containing 5, 10 and 40mM imidazole respectively. 40-1, 40-2, 40-3 = successive fractions of the 40mM wash. Samples were separated on a 10% SDS polyacrylamide gel. Lanes -/+ IPTG contain extracts from 60µl of cells. L/FT contain extracts from 100µl of cells. 5, 10 and 40mM imidazole samples contain extracts from 1ml of cells.

B The 80mM and 100mM imidazole fractions of the nickel agarose column separated on a 10% SDS polyacrylamide gel. -/+ IPTG fractions as in A. 80 and 100mM imidazole lanes contain extracts from 1ml of cells. As for the 40mM wash, three successive fractions (1, 2, 3) were collected during the 80 and 100mM washes.


40
IPTG
3
2
1
10
5
FT
L
+
-
180
116
84
58
48.5
36.5
26.6

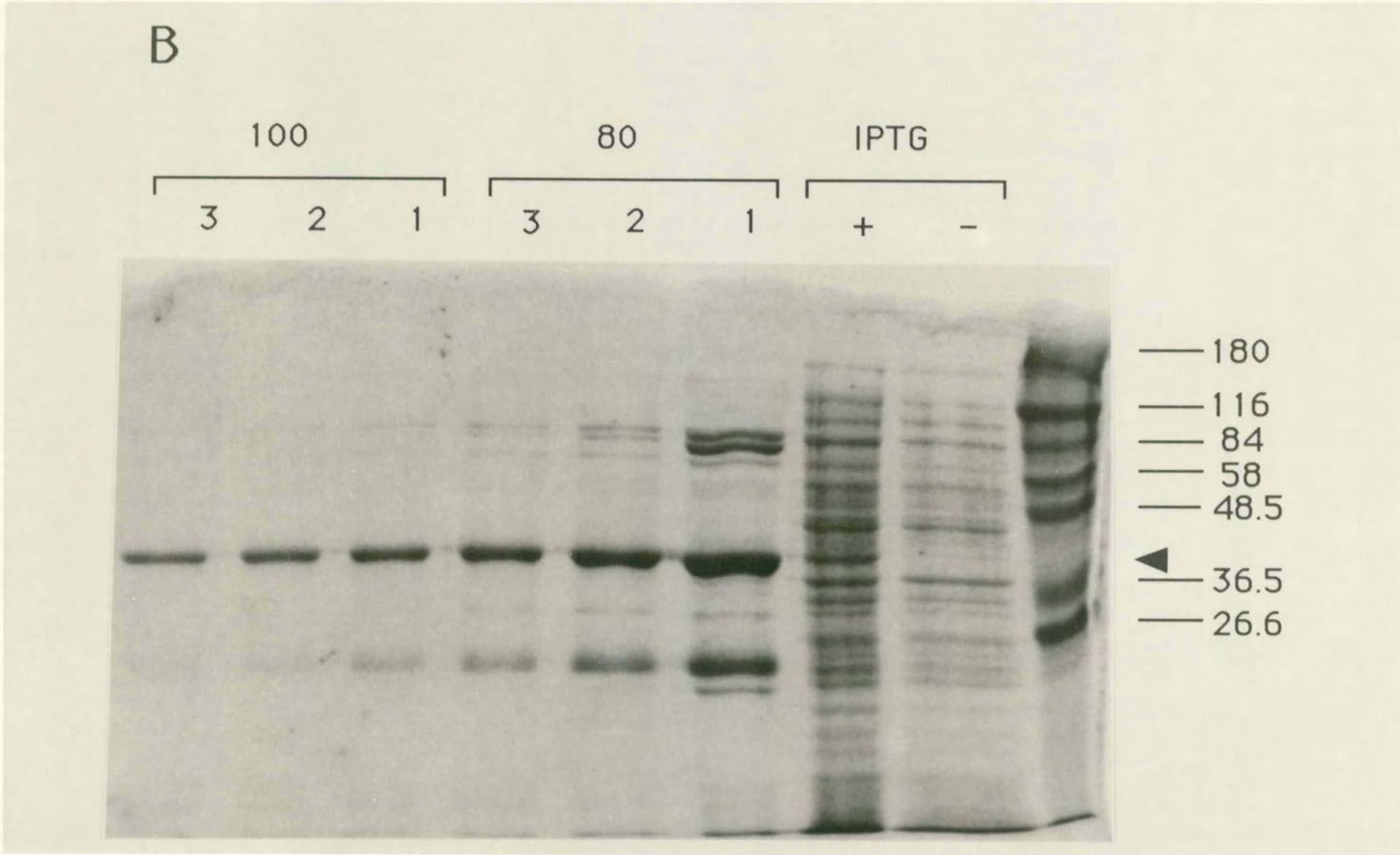
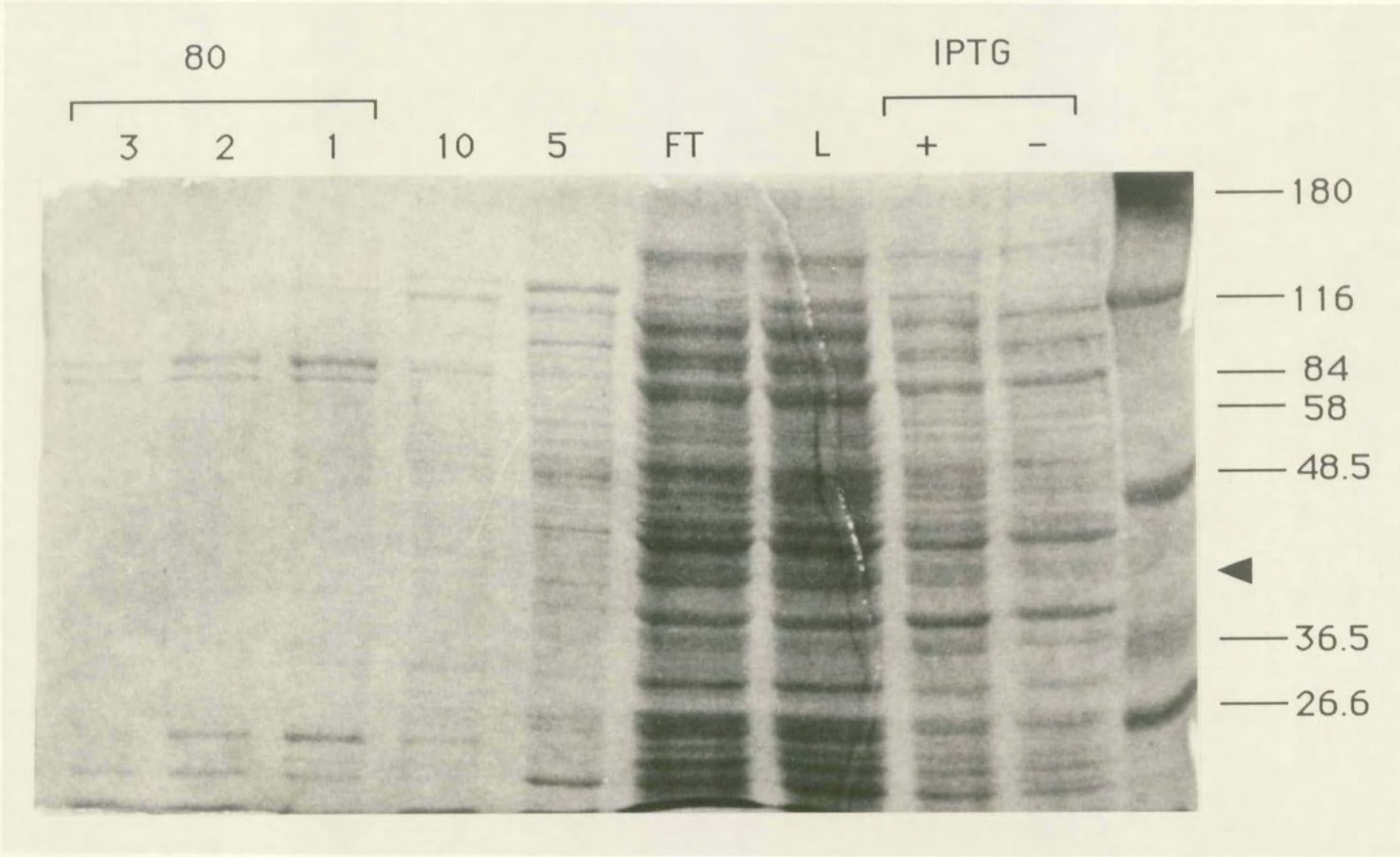

B
100
80
IPTG
3
2
1
3
2
1
+
-
180
116
84
58
48.5
36.5
26.6

Figure 5.6 Nickel ion affinity chromatography of bacterial protein extracts made from cells carrying the pet6H vector.

Experimental procedures were exactly as described in Figure 5.5 except that cells were carrying the pet6H vector. -/+ IPTG lanes contain extracts from 60µl of cells. L/FL lanes contain extracts from 100µl of cells and 5, 10 and 80mM imidazole lanes contain extracts from 1ml of cells.


80
IPTG
3
2
1
10
5
FT
L
+
-
180
116
84
58
48.5
36.5
26.6

5.4 *In vitro* methyltransferase assays

To test whether the cnd1 protein could act as a DNA methyltransferase *in vitro*, the purified protein was incubated with a variety of substrates in the presence of ^{3}H-Adomet (section 2.8 iii). DNA methyltransferase activity was measured by the incorporation of the tritiated methyl groups into the substrates. As a positive control, the bacterial m^{5}C-MTase M.MspI was used. This enzyme recognizes the sequence CCGG and methylates the external cytosine. For assays involving substrates containing repeated CpG dinucleotides, the bacterial m^{5}C-MTase SssI was used. This enzyme methylates cytosines in the context of CpG dinucleotides. These *in vitro* assays rely on the assumption that the substrates that are used contain enough copies of the DNA recognition sequence of the putative cnd1 m^{5}C-MTase to give detectable methyl transfer.

The substrates included total genomic DNA from *S. pombe* (from both wild type and *cnd1-1* strains), bacteriophage lambda, *E.coli* and *Drosophila melanogaster*. Also used were short pieces of double stranded DNA containing repeated CpG dinucleotides. In addition, plasmids containing specific *S. pombe* sequences were used including the dgdh repeats which are located in the centromere (Murakami *et al.*, 1991), the telomeric repeats and telomere-associated sequences (Sugawara and Szostak, 1986) and a short piece of DNA containing the region of the *mat1* locus that acquires a double strand break during the process of mating type switching (details in Appendix A). The rationale for choosing the *mat1* locus deserves special mention. The double strand break in the DNA initiates the process of switching and whether the DNA is cut or not is thought to depend on a

signal in the form of an epigenetic modifiction of the DNA in this region (Beach, 1983). The nature of the signal is unknown but as DNA methylation has the characteristics required for epigenetic modification, it has been proposed as a suitable candidate (Klar, 1987). Thus I was especially interested to test this region as a substrate for the putative cnd1 m^5C-MTase.

In addition to the above DNA substrates, total RNA from fission yeast, both wild type and the *cnd1-1* mutant, was used. Methylation of cytosine at the 5-position has been found in ribosomal RNA in organisms ranging from bacteria to man (Noller, 1984; Maden, 1988), but the cloning of the genes responsible for this modification has not been reported, therefore the possibility that cnd1 could use RNA as its substrate was explored.

It was not possible to detect transfer of methyl groups by cnd1 to any of the substrates used (Table 5.1). By contrast, M.MspI and, where appropriate, SssI were active under the same assay conditions. In addition to the standard assay (section 2.8 iii), a number of variant conditions were tested without effect. These included a range of incubation temperatures from 25ºC - 37ºC, a range of Mg^{2+} concentrations from 0-10mM and a range of ^3H-AdoMet concentrations from 6.7µM to 80µM. *S. pombe* protein extracts were also added to the incubation mix in case the cnd1 protein required association with or modification by certain fission yeast proteins before it could become an active methyltransferase.

5.5 Specific antisera against the cnd1 protein

Specific antisera against the cnd1 protein were raised in rabbits. The preparation of the protein for injection is shown in Figure 5.7 and the details and time course of the doses are described in section 2.8 iv. Antiserum from

one particular rabbit 722, showed a high titre in western blots of purified cnd1. This antiserum was also able to recognize cnd1 in protein extracts made from a fission yeast strain over-expressing the *cnd1*$^{+}$ gene under the control of the nmt1 promoter in the plasmid pREP1 (Figure 5.8; sections 6.2 i; 2.2 v). The over-expressed cnd1 protein in this blot appeared to be the same molecular weight as the purified cnd1 indicating that it had not been grossly modified in the yeast cells. The antisera do not, however, recognize an endogenous protein in wild type cells (Figure 5.8). This could be attributable to the quality of the antibodies or may be a consequence of low or absent expresssion of the cnd1 protein in fission yeast cells.

5.6 Conclusions

The cnd1 protien was histidine-tagged at its N-terminus and purified from bacterial cells by nickel ion affinity chromatography. *In vitro* methyltransferase assays were performed using the purified protein but it was not possible to detect transfer of ^{3}H-methyl groups by cnd1 to any of the substrates used. One possible reason for the apparent lack of activity could be that the protein has to be modified in some way to become active. Alternatively, the tritium transfer assay may not be a sensitive enough method to detect small amounts of m^{5}C-MTase activity as would occur if, for example, the recognition sequence for cnd1 was more than 7bp long. A further explanation for the lack of transfer is that cnd1 is not a functional m^{5}C-MTase. These possibilities will be discussed further in chapter 8.

Specific antisera were generated against the purified cnd1 protein. The antibodies were able to detect cnd1 when over-expressed from a heterologous promoter but did not identify the endogenous protein in wild type cells.

Table 5.1 A. Methyl group incorporation by cnd1 and MspI into genomic DNA and RNA substrates.

Methyltransferase	λ	*E. coli*	WT[b]	*cnd1-1*[b]	*Drosophila*	WT [c]	*cnd1-1* [c]
none[a]	38	51	50	47	111	39	62
cnd1	30	55	30	39	86	55	36
MspI	30084	31900	2140	2280	n.d.	n.d.	n.d.
	10μg	30μg	1μg	1μg	100ng	10μg	10μg

"n.d." denotes not determined. Methylation was determined by the transfer of the radiolabelled methyl group from [^{3}H]-SAM into DNA. Values given are in c.p.m. The amount of each substrate used is indicated at the bottom of each column. The background was typically 20-30cpm and this was not subtracted from the figures shown in tables 5.1A and 5.1B. Each assay was performed three times and the average value was taken. The values for samples containing no substrate were also typically between 20 and 30 cpm for both cnd1 and MspI. The lambda DNA that was used had been prepared from dcm$^-$ dam$^-$ *E. coli* cells and therefore is not methylated at any sites.

[a] An extract from cells carrying plasmid pET6H (the expression vector alone) was passed over a nickel agarose column and the 100mM imidazole fraction was used as a negative control in these assays. λ and *E. coli* refers to genomic DNA from these organisms.

[b] refers to genomic DNA from *S. pombe* (either wild type (WT) or from the *cnd1-1* mutant as indicated.

[c] refers to total RNA from *S. pombe*, either from wild type or the *cnd 1-1* mutant as indicated.

Table 5.1 B. Methyl group incorporation by cnd1 and MspI into other DNA substrates.

Methyltransferase	centromere[d]	telomere[e]	GAC/GTC[f]	GAC/GTM[f]	poly dG:dC[g]	poly dI:dC[g]	DSB[h]
none[a]	37	29	59	59	48	76	56
cnd1	29	40	52	53	41	56	49
MspI	16015	1084	n.d.	n.d.	8015[i]	17615[i]	n.d.
	20µg	10µg	10µg	10µg	40µg	40µg	10µg

[d] *S. pombe* centromeric dgdh repeats (Murakami *et al.*, 1991).

[e] *S. pombe* telomeric DNA and the telomere associated sequences (Sugawara and Szostak, 1986).

[f] GAC/GTC contains 12 copies of GAC flanked at each end by *Sau3AI* sites to give a 40mer containing 12 CG pairs. GAC/GTM is identical except that one strand contains methylated cytosines and is therefore a hemimethylated substrate.

[g] poly dG: dC and poly dI: dC were obtained from Pharmacia and are polymers consisting of repeated GC or IC dinucleotides respectively.

[h] DSB refers to the Double strand break region involved in *S. pombe* mating type switching (see Appendix A).

[i] SssI was used in these assays instead of MspI.

Figure 5.7 Purification of cnd1 protein for the preparation of polyclonal antisera

The cnd1 protein was purified by nickel agarose chromatography in high salt buffer (500mM NaCl) as described (section 2.8 ii). The 80-1 and 80-2 fractions were dialyzed against a low salt buffer and the resulting protein sample centrifuged at 14K in a microcentrifuge for 10 minutes.
Total=sample before dialysis, S=supernatant and P=pellet fractions after the centrifugation step. Known amounts of bovine serum albumin were run in order to estimate the quantities of the cnd1 protein. Samples were separated on a 10% SDS polyacrylamide gel.

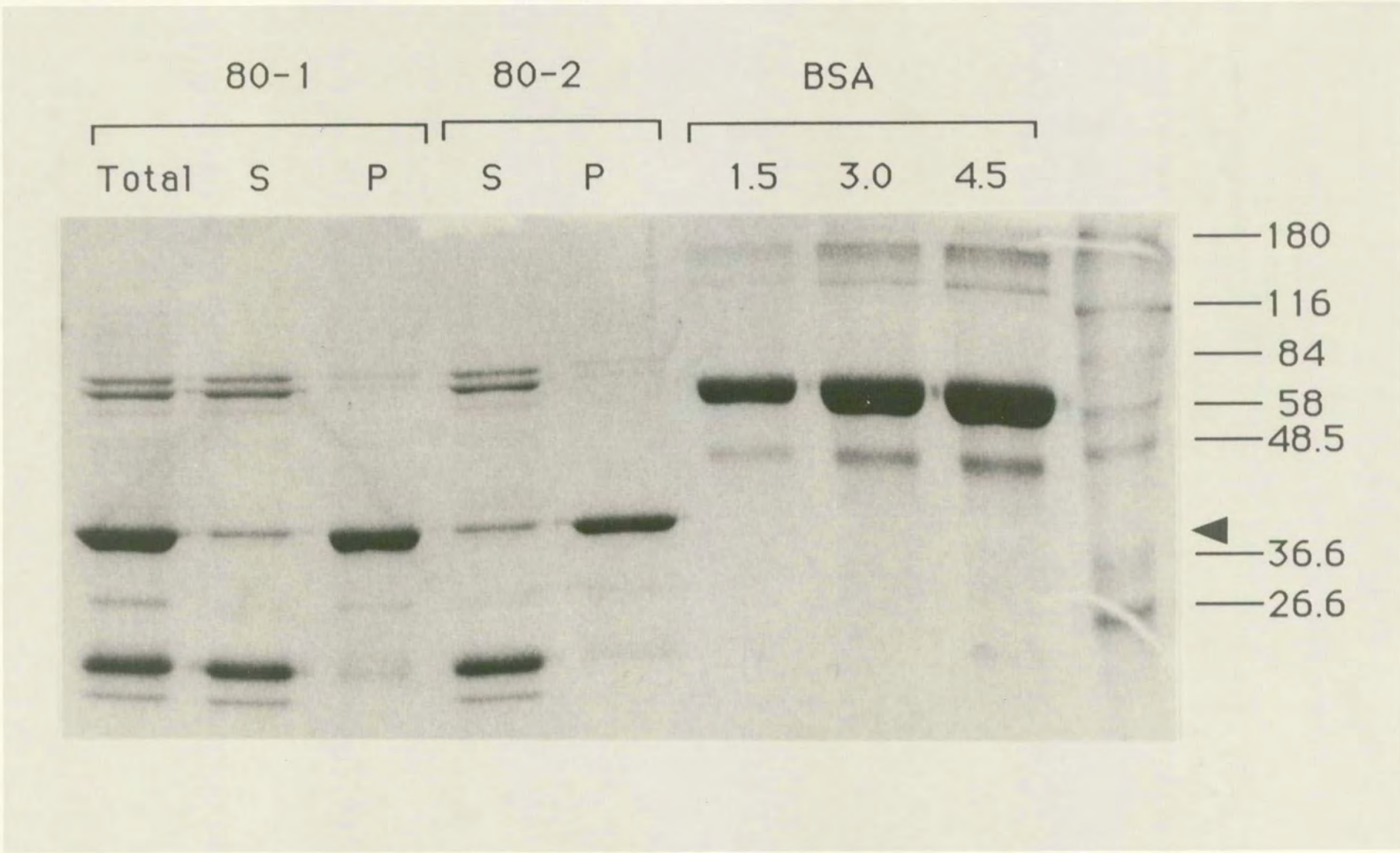

80-1
80-2
BSA
Total
S
P
S
P
1.5
3.0
4.5
180
116
84
58
48.5
36.6
26.6

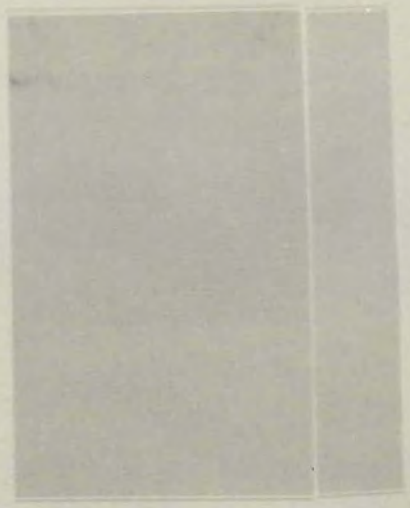
Figure 5.8 Western blots of *S. pombe* protein extracts using polyclonal antisera raised against the cnd1 protein

Protein extracts were separated by polyacrylamide gel electrophoresis and then transferred to nitrocellulose membranes. Membranes were incubated with either serum (diluted 1 in 15000) as indicated; **A**, preimmune; **B**, antisera raised against the cnd1 protein (for details of experimental procedures see section 2.8 v). Lane 1: 5ng of cnd1 protein purified by nickel ion affinity chromotography;
lane 2: 50µg of extracts made from *S. pombe* cells containing pREP1cnd1;
lane 3: 50µg of extracts made from *S. pombe* cells containing pREP1;
lane 4: 50µg of extracts made from *S. pombe* wild type cells. The positions of the molecular weight standards that were transferred from the gel onto the nitrocellulose membrane are indicated in kilodaltons. Films in both panels were exposed to ECL film for 1 minute.

A

Preimmune

1 2 3 4

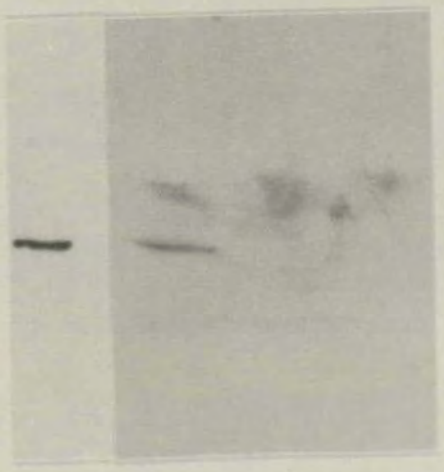


97.4
66
45
31
21.5

B

Ab 722

1 2 3 4



97.4
66
45
31
21.5

Chapter 6: The *S. pombe* m^5C-MTase homologue is not the *cnd1*$^+$ gene

6.1 Introduction

One possible explanation for the failure to detect methyltransferase activity with the recombinant histidine-tagged cnd1 protein (Chapter 5) was that the protein had to be activated by modification processes *in vivo*. In order to allow for possible modifications to occur, I decided that the most suitable course of action would be to over-express the cDNA in fission yeast and then purify the resulting cnd1 protein before attempting further methyltransferase assays. To facilitate purification of the native protein, it was planned to fuse it in frame with an epitope such as an influenza antigen (Kolodziej and Young, 1991). The tagged protein could then be immuno-precipitated from *S. pombe* protein extracts using monoclonal antibodies against the tag. The resulting cnd1 protein could then be used in methyltransferase assays as described in chapter 5.

Before proceeding with this scheme, however, it was first necessary to determine whether cells over-expressing the *cnd1*$^+$ gene were viable. As a control for this analysis, the construct over-expressing the *cnd1*$^+$ cDNA was also transformed into the *cnd1-1* mutant cells and surprisingly was unable to rescue the temperature-sensitive lethality associated with this mutation. The resulting investigation of this anomaly is described in this chapter.

6.2 Over-expression of the *cnd1+* gene in *S. pombe*

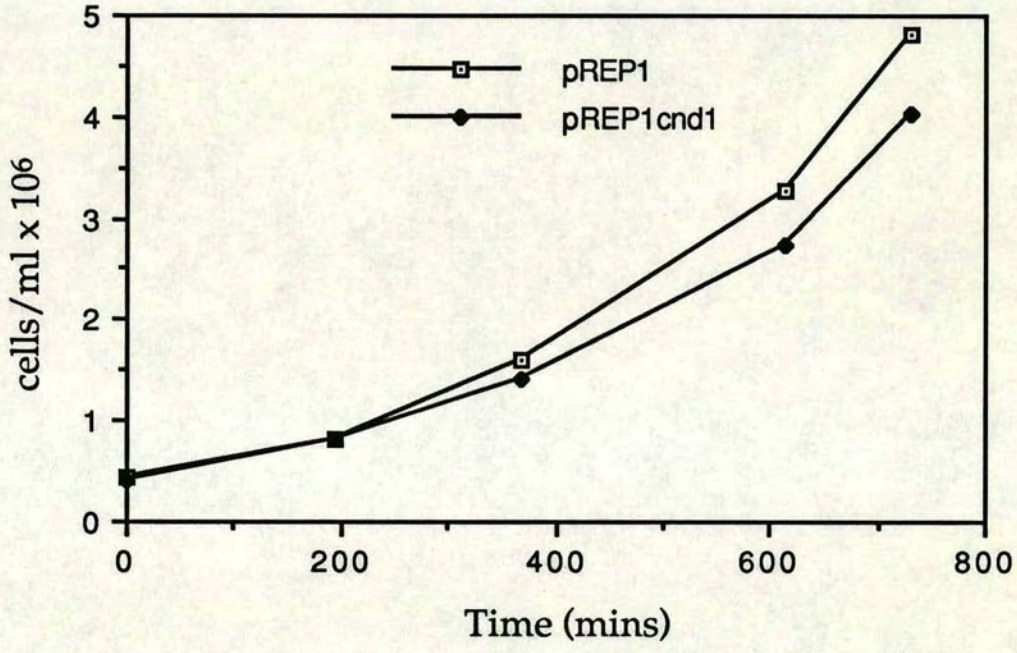
i Over-expression of *cnd1+* is not lethal in wild type cells

The *cnd1* cDNA was first subcloned into the vector pREP1 (Maundrell, 1990; Figure 2.3, section 2.2 v) which allows substantial over-expression of genes under the control of the nmt1 promoter (Table 2.1). Expression from nmt1 is induced in the absence and repressed in the presence of thiamine. The *cnd1+* cDNA was amplified by PCR from the C145 plasmid which contains the 1.1kb *cnd1+* cDNA subcloned into the *HindIII* site in pTZ19R. Oligonucleotide primers were designed such that the start ATG of the *cnd1+* gene was incorporated into an *NdeI* site. A *BamHI* site was positioned 7 nucleotides downstream of the TAA stop codon (details of the oligonucleotides and the PCR reaction can be found in section iii of Appendix B). The cDNA was then subcloned into the *NdeI* and *BamHI* sites of the pREP1 vector to give the plasmid pREP1cnd1.

The wild type strain *h⁻ leu1-32* was transformed with pREP1 and pREP1cnd1. The pREP1 plasmid carries the *LEU2* marker and therefore transformants were selected by their ability to grow on leucine-deficient minimal medium. The medium contained no thiamine to ensure full induction of *cnd1+* expression. Many transformants were obtained indicating that over-expression of the *cnd1+* gene is not lethal. The cells appeared indistinguishable from wild type under the light microscope. In addition, the generation time of cells transformed with pREP1cnd1 was compared to a strain carrying the pREP1 vector alone. Both strains had a generation time of 225 minutes at 27ºC (Figure 6.1).

Figure 6.1 The over-expression of the *cnd1*$^+$ gene does not affect the growth rate of wild type *S. pombe* cells

S. pombe leu1-32 cells containing either pREP1cnd1 or pREP1 were grown in minimal medium without thiamine at 27ºC for 12 hours to a cell number of 1×10^7 cells ml^{-1}. The cells were then diluted into fresh medium to a concentration of approximately 4×10^5 cells ml^{-1} and allowed to continue growing. Samples of culture (100µl) were removed at regular intervals and the number of cells were counted using a Coulter counter. Duplicate samples were taken at each time point and the average value taken to calculate the number of cells ml^{-1}. The concentrations of the cultures upon dilution into fresh medium (time = 0) were 4.50×10^5 and 3.98×10^5 cells ml^{-1} for cells transformed with pREP1 and pREPcnd1 respectively. This accounts for the slight difference between the two curves, despite both cultures having identical generation times.


pREP1
pREP1cnd1
cells/ml x 10⁶
0
1
2
3
4
5
0
200
400
600
800
Time (mins)

ii Over-expression of the *cnd1+* gene does not rescue the *cnd1-1* mutant

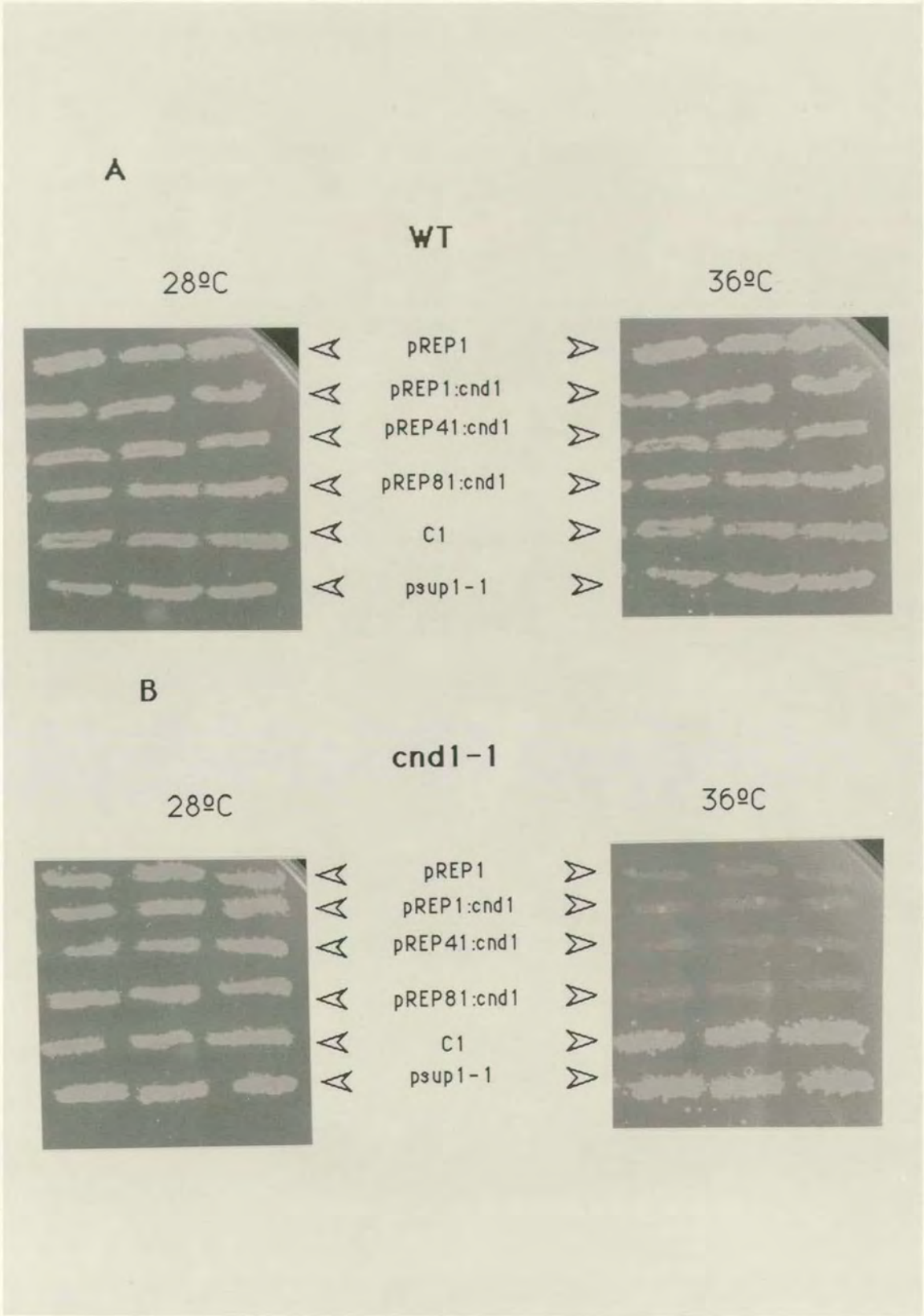
In addition to transforming wild type cells, pREP1cnd1 was also transformed into *cnd1-1 leu1-32* mutant cells, to check for expression of the *cnd1+* gene. If the temperature-sensitive mutation was rescued by pREP1cnd1, this would indicate that the *cnd1+* gene was being expressed from the nmt1 promoter. Surprisingly, pREP1cnd1 failed to rescue the lethality of the *cnd1-1* mutation at 36°C (Figure 6.2). There were a number of hypotheses to account for this finding and the ways in which they were tested are discussed below.

The lack of complementation could have been a result of the cnd1 protein not being synthesized. For example, if the *cnd1* mRNA was unstable when induced from a heterologous promoter, it might have been degraded before translation could occur. However, as described in section 5.5, polyclonal antisera raised against cnd1 detected a protein of the correct size in extracts made from *S. pombe* cells containing pREP1cnd1, indicating that the cnd1 protein was synthesized under the control of the nmt1 promoter (Figure 5.8).

An alternative hypothesis to explain the lack of rescue by pREP1cnd1 was that the *cnd1+* cDNA had acquired a point mutation which altered the protein sequence. Although the fidelity of Taq polymerase is generally high, occasional errors are made during PCR reactions. As the *cnd1+* cDNA was amplified by PCR in order to introduce suitable restriction sites at the 5' and 3' termini, it was possible that a point mutation had altered an essential codon in the *cnd1+* coding sequence. To test whether this was the case, the *cnd1* sequence in pREP1cnd1 was sequenced on one strand using the

Figure 6.2 Over-expression of the *cnd1*$^+$ gene does not rescue the *cnd1-1* mutation

Fission yeast cells were transformed with various plasmids as described in the text (section 6.2 ii). Transformants were patched onto minimal medium plates (containing no thiamine) and allowed to grow for 24 hours. The plates were incubated at either the permissive or restrictive temperatures for the *cnd1-1* mutation (28ºC or 36ºC respectively) as indicated. The strains transformed were the wild type (WT) strain *h*$^-$ *leu1-32* (panel A) and the temperature sensitive mutant strain *h*$^-$ *leu1-32 cnd1-1* (panel B). The psup1-1 plasmid contains the genomic clone of *scd1*$^+$ (the multicopy suppressor of *cnd1-1*).


A
WT
28ºC
36ºC
pREP1
pREP1:cnd1
pREP41:cnd1
pREP81:cnd1
C1
psup1-1
B
cnd1-1
28ºC
36ºC
pREP1
pREP1:cnd1
pREP41:cnd1
pREP81:cnd1
C1
psup1-1

oligonucleotides cndA-E and the 861 oligonucleotide as described in Appendix B. No mutations were found in the *cnd1* sequence disproving this explanation of the lack of complementation.

Although over-expression of the *cnd1*$^+$ gene did not affect wild type cells, it was possible that a combination of high levels of the cnd1 protein along with the defects associated with the *cnd1-1* mutation might prove lethal. Therefore, the *cnd1*$^+$ cDNA was subcloned into the vectors pREP41 and pREP81 which contain the nmt1 promoter with mutations in its TATA box which result in reduced levels of expression of genes under its control (Basi *et al.*, 1993, section 2.2 v; Table 2.1). The pREP41cnd1 and pREP81cnd1 constructs were also unable to rescue the *cnd1-1* mutation (Figure 6.2).

6.3 The *cnd1-1* mutant has no mutation in the m^{5}C-MTase homologue gene

As there appeared to be no obvious reason why pREP1cnd1 did not rescue the *cnd1-1* mutation, it seemed increasingly likely that the m^{5}C-MTase homologue was not the real *cnd1*$^+$ gene. To investigate this possibility, genomic DNA was made from the *cnd1-1* mutant and the region containing the two m^{5}C-MTase ORFs amplified by PCR using the oligonucleotides Ndecnd5 and cnd13. The resulting DNA products from three PCR reactions were subcloned into the pREP1 vector and two clones from each reaction were sequenced using the oligonucleotides cndA-E and the 861 oligonucleotide (Appendix B). This analysis was carried out using an ABI automated sequencing machine with the assistance of N. Tountas. Analysis of the sequence data revealed that there was no mutation in either of the two ORFs corresponding to the m^{5}C-MTase homologue. The inescapable conclusion was that the *cnd1*$^+$ gene, as defined by the *cnd1-1*

mutation (Bartlett, 1991) was not the m5C-MTase homologue. Consequently, the gene encoding the m^5C-MTase homologue was renamed *pmt1*$^+$ (pombe methyltransferase).

6.4 The identity of the *cnd1*$^+$ gene

Although the *cnd1-1* phenotype was no longer associated with the m^5C-MTase homologue, I felt that it was important to identify the real *cnd1*$^+$ gene for several reasons. Firstly, the consequences of the *cnd1-1* mutation clearly implied that the *cnd1*$^+$ gene function was important. In particular, the specific phenotypes associated with the mutation seemed to suggest that this gene was involved not only in nuclear division but also in the regulation of the switch between the mitotic and meiotic cell cycles. Secondly, as the suppressor of the *cnd1-1* mutant, *scd1*$^+$, had been sequenced, it seemed pertinent to establish the correct identity of the mutated gene.

The C1 plasmid containing the *cnd1*$^+$ genomic clone was described as having one 2.4kb insert (Bartlett, 1991). Examination of the sequence of this insert revealed that there was only one major ORF corresponding to the pmt1 protein (plus the smaller pmt1 ORF located upstream). Hence the location of the *cnd1*$^+$ gene did not seem obvious. To try and narrow down the region containing the rescuing activity, I decided to subclone parts of the 2.4kb insert into pDB262 and check for rescue of the mutation. Thus the C1 plasmid was obtained from R. Bartlett and P. Nurse. It was already clear that this plasmid contained the real *cnd1*$^+$ gene as this had been demonstrated by integration of the plasmid at the *cnd1-1* locus (Bartlett, 1991; see Introduction section 1.7).

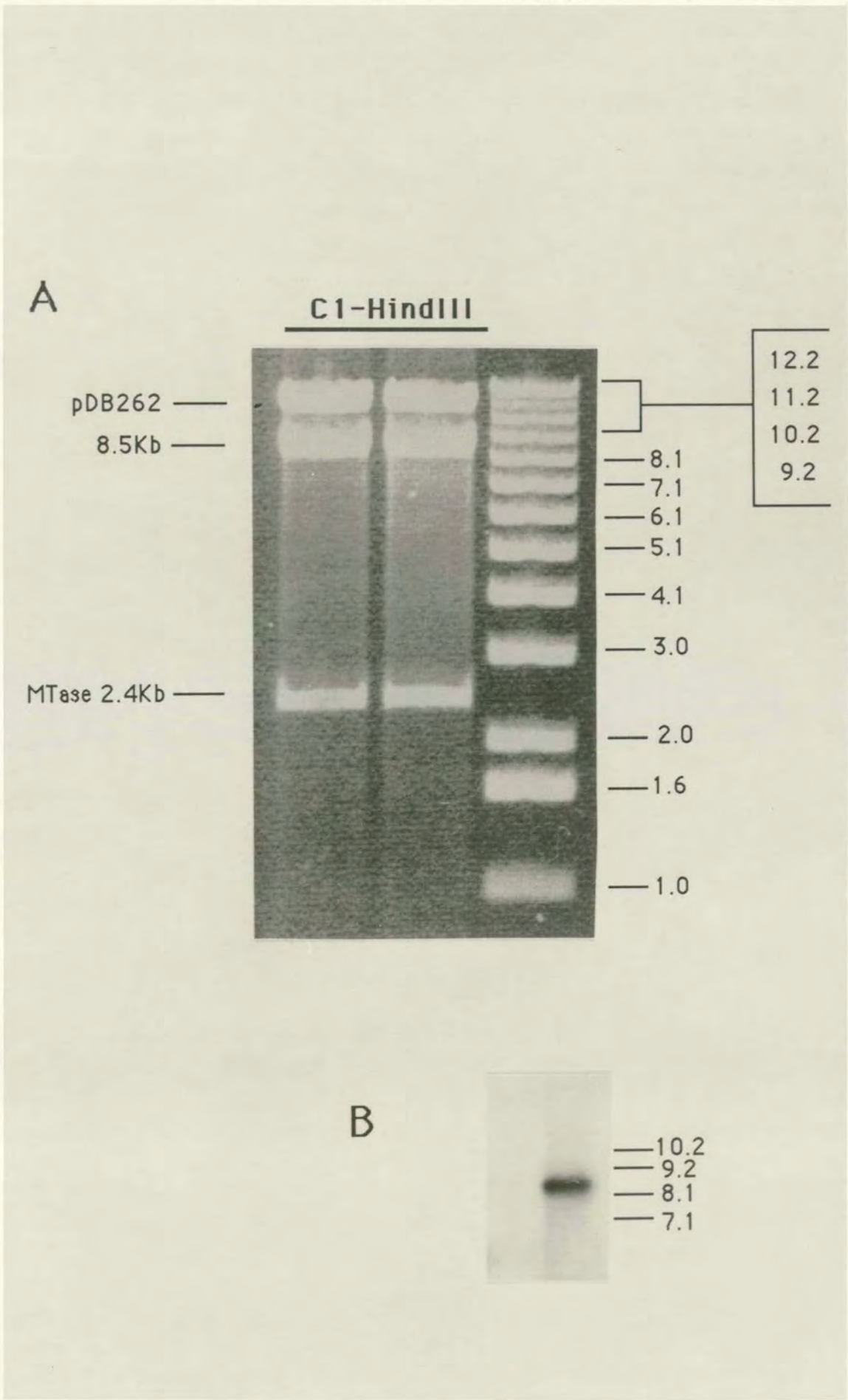
After transformation into *E. coli*, C1 plasmid was extracted and digested with *HindIII* to check that no rearrangements of the DNA had

Figure 6.3 A The C1 plasmid contains two *HindIII* inserts

C1 plasmid (1ng) was transformed into *E. coli* cells and plasmid DNA prepared from the resulting transformants. DNA obtained from two separate transformants was digested with *HindIII* for 3 hours at 37ºC and then resolved on an 0.8% agarose gel (each lane contains 0.5µg of DNA). Molecular weight standards are indicated in kilobases. The expected positions of the 2.4kb fragment containing the m^5C-MTase homologue and the 10kb pDB262 vector band are indicated along with the unexpected 8.5kb band.

B Southern blot analysis to show that the 8.5kb band in C1 is derived from fission yeast DNA

Southern blot of 0.5µg of *S. pombe* DNA digested with *HindIII* and hybridized to a probe made from the 8.5kb fragment in the C1 plasmid. Film was exposed overnight at -70ºC. Molecular weight standards are indicated in kilobases.


A
C1-HindIII
pDB262
8.5Kb
MTase 2.4Kb
12.2
11.2
10.2
9.2
8.1
7.1
6.1
5.1
4.1
3.0
2.0
1.6
1.0
B
10.2
9.2
8.1
7.1

occurred during propagation in the bacterial cells. Surprisingly, two insertswere found in the C1 plasmid (Figure 6.3A). In addition to the expected 2.4kb fragment, there was also an 8.5kb insert. The digest did not appear to be partial and the appearance of the 8.5kb band was reproducible. The 8.5kb insert hybridized to a band of identical size in a Southern blot of *S. pombe* DNA digested with *HindIII* indicating that this insert was derived from fission yeast DNA and was not due to a rearrangement of the vector (Figure 6.3B).

It seemed possible that this hitherto undetected fragment, contained the real *cnd1*$^+$ gene. Southern blot analysis indicated that the 2.4kb and 8.5kb inserts were also linked in the genome as they recognized bands of identical size in *EcoRI* and *BamHI* digests of *S. pombe* DNA (data not shown). Although the library from which the C1 plasmid was isolated was constructed using a complete *HindIII* digest of fission yeast genomic DNA, the site between the 8.5kb and 2.4kb fragments presumably escaped cleavage thereby accounting for the presence of the two inserts in C1.

To determine whether the *cnd1*$^+$ gene was located within the 8.5kb insert, this fragment was subcloned into the pDB262 vector and the resulting pDB262:8.5 plasmid transformed into the *cnd1-1* mutant. The transformants were able to grow at 36ºC indicating that the 8.5kb insert did indeed contain the real *cnd1*$^+$ gene (Figure 6.4). The rescuing activity was narrowed down to a 3.5kb *HindIII-BamHI* sub-fragment which was subcloned into pBluescript KS$^-$. A restriction map of the insert was generated and a number of sub-fragments of the 3.5kb insert subcloned into pBluescript KS$^-$ for sequence analysis (Figure 6.5). The identity of the *cnd1*$^+$ gene was not revealed by sequencing these subclones, however, but came instead by a rather different route.

Figure 6.4 The 8.5kb fragment from the C1 plasmid is able to rescue the *cnd1-1* mutant

The fission yeast strain *cnd1-1 leu1-32* was transformed with the following constructs as indicated; the pDB262 vector **(pDB262)**; the original C1 plasmid containing both the 2.4kb and 8.5kb *HindIII* fragments **(2.4 + 8.5)**; the 8.5kb fragment from C1 subcloned into pDB262 **(8.5)**; the 2.4kb m^5C-MTase homologue fragment subcloned into pDB262 **(2.4)**. The resulting transformants were streaked to single colonies on minimal medium plates and incubated for 3 days at 28ºC (panel A) or 36ºC for 2 days (panel B).

A 28ºC

pDB262

2.4

2.4 + 8.5

8.5

B 36ºC

pDB262

2.4

2.4 + 8.5

8.5

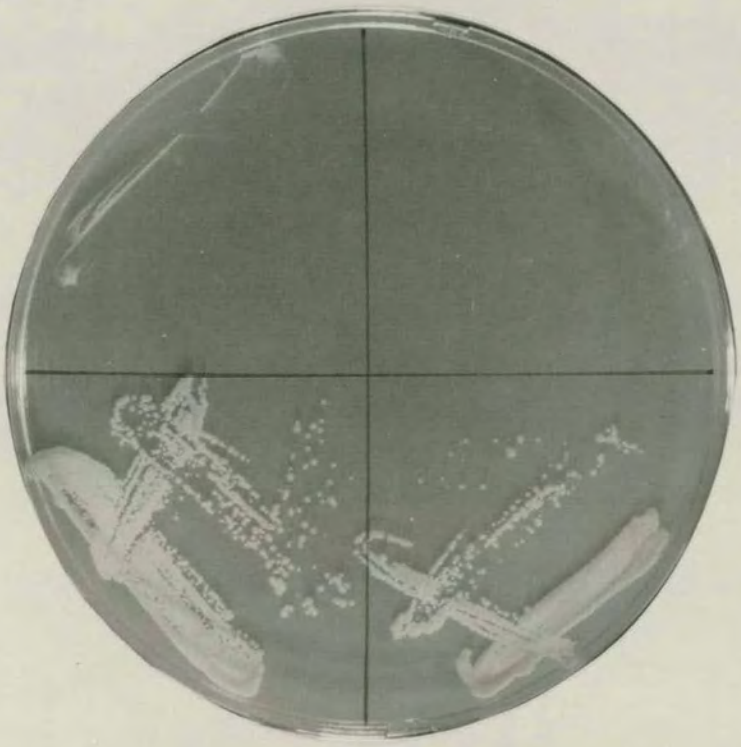


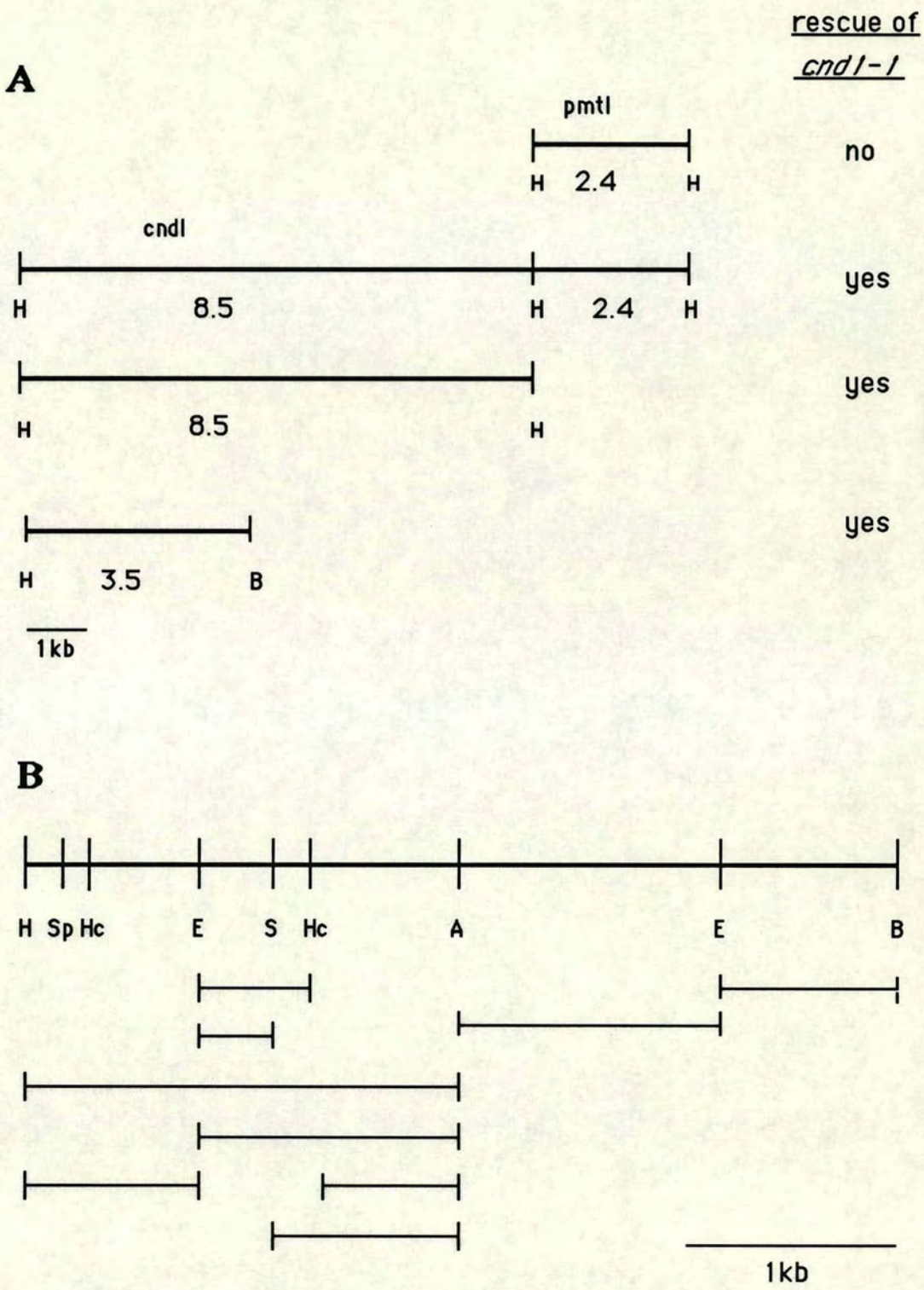
Figure 6.5

A **A summary of the inserts in the C1 plasmid which are able to rescue the *cnd1-1* mutant.**

The 2.4kb *pmt1*$^+$ and the 8.5kb *pat1*$^+$ genomic clones are indicated. H = *HindIII*, B = *BamHI*. The 3.5kb *HindIII-BamHI* fragment that is able to rescue *cnd1-1* is shown.

B **Restriction map of the 3.5kb *HindIII-BamHI* fragment containing the *cnd1*$^+$ gene**

H = *HindIII*, Sp = *SphI*, E = *EcoRI*, S = *SacI*, Hc = *HincII*, A = *Asp718*, B = *BamHI*. Below the map are the sub-fragments that were subcloned into pBluescript KS$^-$ for sequencing.


rescue of
cnd1-1
A
pmtl
H 2.4 H
no
cndl
H 8.5 H 2.4 H
yes
H 8.5 H
yes
H 3.5 B
yes
1kb
B
H Sp Hc E S Hc A E B
1kb

At the same time as these experiments were carried out, the genomic location of *pmt1*$^+$ (the m^{5}C-MTase homologue) was determined byhybridization of its cDNA to *S. pombe* cosmid and bacteriophage P1 libraries. The details of this investigation are described in chapter 7. As *pmt1*$^+$ and *cnd1*$^+$ were presumed to be closely linked in the genome, this analysis was also expected to reveal the location of the *cnd1*$^+$ gene.

The *pmt1*$^+$ gene was found to be located on chromosome II, next to the centromere, within cosmid 10G12C on the NotI-B fragment between the genetic markers *pat1* and *cen2*. The *pat1*$^+$ gene encodes a protein kinase that plays a major role in switching from vegetative growth to sexual development. Its inactivation results in unconditional growth arrest and subsequent meiosis regardless of DNA content, nutritional conditions and the availability of cells of the opposite mating type (Nurse, 1985; Iino and Yamamoto, 1985 and McLeod and Beach, 1986).

The phenotypes of *cnd1-1*, namely entry into meiosis in rich medium and the increase in DNA content without nuclear division, appeared very similar to those associated with mutations in *pat1*. A restriction map of the 3.5kb *pat1*$^+$ genomic clone (McLeod and Beach, 1986) was compared to that of the 3.5kb fragment known to contain the *cnd1*$^+$ gene. The two maps were identical, strongly implying that *cnd1*$^+$ was the *pat1*$^+$ gene. Further confirmation was obtained by sequencing 400 base pairs of DNA at each end of the 3.5kb *HindIII-BamHI* fragment (this sequencing was performed by V. Clarke). The sequence obtained was a perfect match to the ends of the *pat1*$^+$ genomic clone (McLeod and Beach, 1986) and so the *cnd1-1* mutant was renamed *pat1-8*. The rescue of this mutant by the 3.5kb *HindIII-BamHI pat1*$^+$ genomic clone is illustrated in Figure 6.6.

Figure 6.6 **The 3.5kb *HindIII-BamHI* fragment containing *pat1*$^+$ is able to rescue the *cnd1-1* mutant**

The *cnd1-1* mutant cells were transformed with the following plasmids as indicated; pIRT2 vector (pIRT2); pIRT2 containing the 3.5kb *HindIII-BamHI pat1*$^+$ genomic clone (pIRT2:pat1), pIRT2 containing the 3kb *EcoRI-SalI scd1*$^+$ genomic clone (pIRT2:scd1). The resulting transformants were streaked to single colonies on minimal medium plates and incubated at 28ºC for 3 days (panel A) or 36ºC for 2 days (panel B).

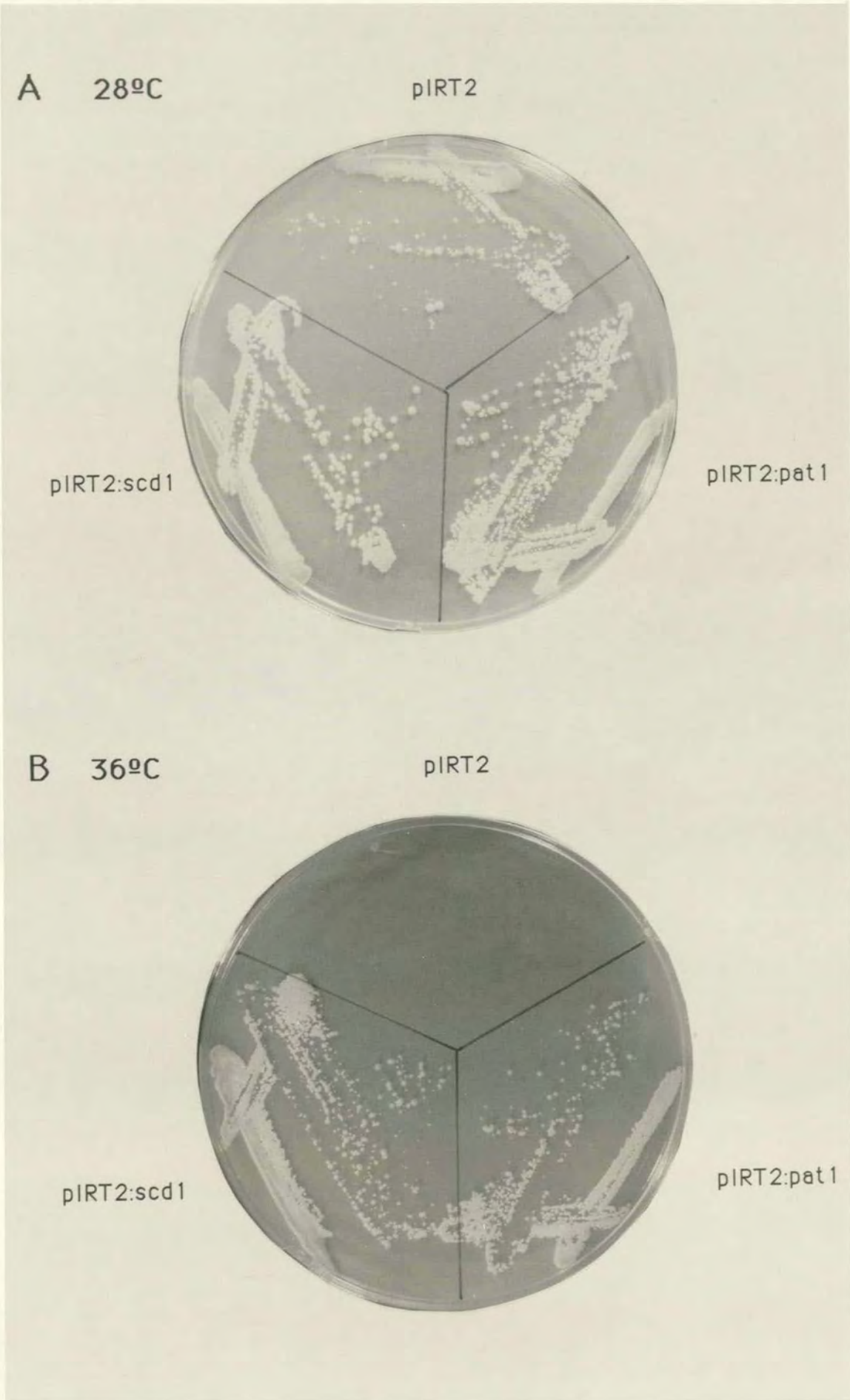
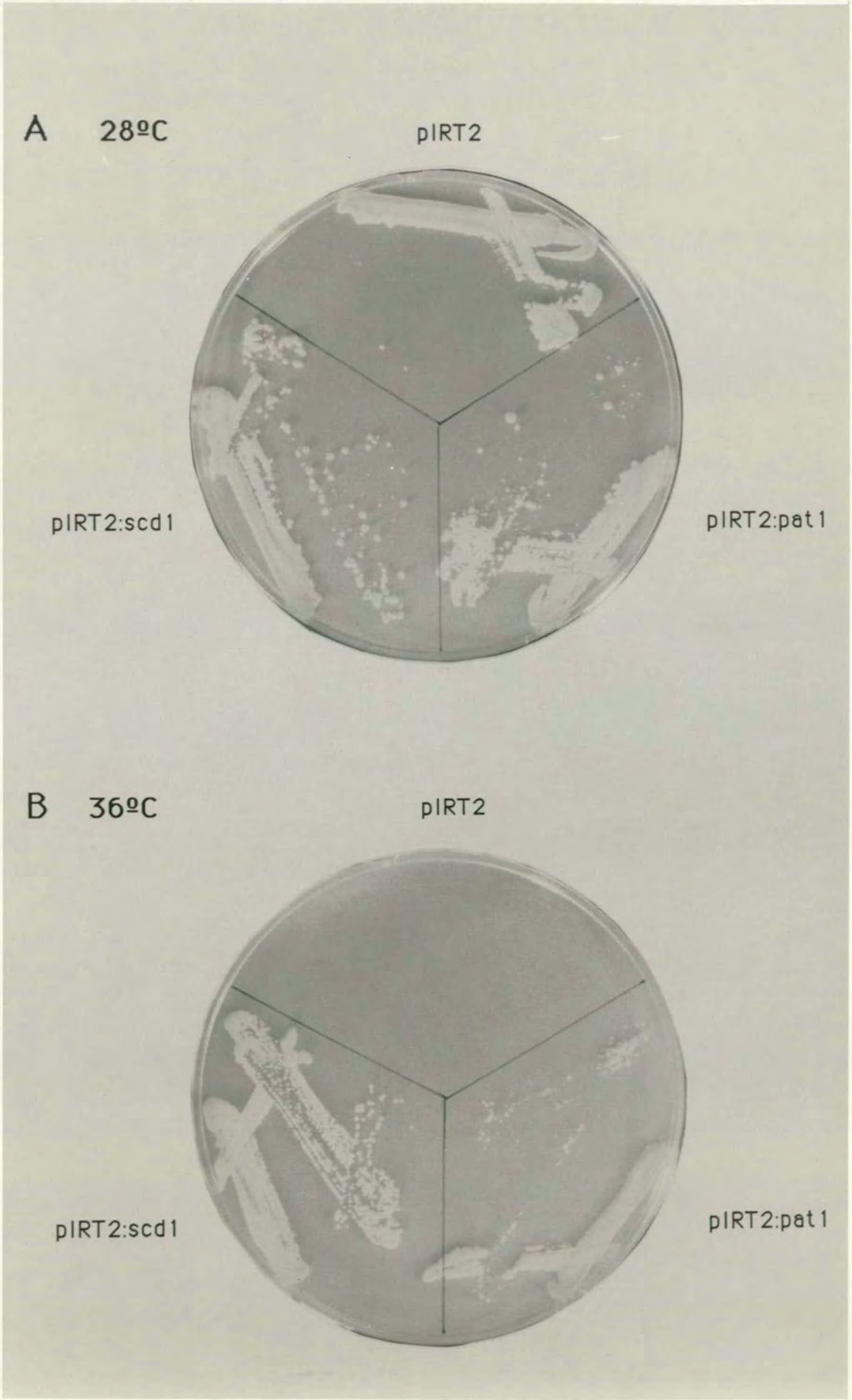

A 28ºC
pIRT2
pIRT2:scd1
pIRT2:pat1
B 36ºC
pIRT2
pIRT2:scd1
pIRT2:pat1

Figure 6.7 The *scd1*$^{+}$ gene is able to rescue a more severe *pat1* mutation

Transformations were carried out exactly as described in Figure 6.6 except that the plasmids were transformed into the *pat1-114* temperature-sensitive mutant.


A 28ºC
pIRT2
pIRT2:scd1
pIRT2:pat1
B 36ºC
pIRT2
pIRT2:scd1
pIRT2:pat1

6.5 *scd1*+ is a suppressor of the *pat1-8* temperature sensitive mutation

With the correct identification of the *cnd1*+ gene, it followed that the *scd1*+ gene (chapter 4) was now known to be a multicopy suppressor of a *pat1* temperature-sensitive mutation. As *pat1*+ was known to play a critical role in the regulation of the meiotic pathway it seemed pertinent to carry out further analysis with the *scd1*+ gene. Most of the analysis of the *pat1*+ gene has been carried out with a particular temperature-sensitive allele *pat1-114* (Iino and Yamamoto, 1985). A comparison by microscopic examination and staining of cells with phloxin B revealed that the phenotype caused by the *pat1-114* mutation is more severe than that associated with *pat1-8*. The *pat1-114* cells growing at 28ºC on rich medium seem to enter meiosis far more readily than *pat1-8* cells and appear less viable at 32ºC (data not shown). The more severe *pat1* allele is also suppressed by over-expression of the *scd1*+ gene (Figure 6.7).

In order to investigate further the functions of *scd1*+, I decided to clone a cDNA and create a null mutant of this gene. Before proceeding with this analysis, however, a restriction map of *scd1*+ was sent to other investigators in the field to ensure that the gene had not already been identified. Unfortunately, it transpired that the gene had been cloned in 1989 by A. Sugimoto and M. Yamamoto and had been named *pac2*+ (M. Yamamoto, pers. comm.). The sequence had neither been published nor submitted to any database. A null mutant of *pac2*+ had already been made. The deletion of the *pac2*+ gene was found to enhance mating indicating that the *pac2*+ function is also closely involved in the regulation of the meiotic pathway. As an extensive investigation had already been carried into the

pac2$^+$ gene function, no further analysis of the *scd1*$^+$/*pac2*$^+$ gene was carried out.

6.6 Analysis of the *pmt1*$^+$ transcript

The finding that the m^{5}C-MTase homologue gene, *pmt1*$^+$, was not associated with the temperature-sensitive phenotype or the derepression of meiosis raised a number of important questions. First of all, was there any phenotype associated with the loss of function of the *pmt1*$^+$ gene? The impetus behind this entire project was provided by the belief that, despite the failure to detect DNA methylation in fission yeast, the function of this m^{5}C-MTase homologue was important due to the phenotype associated with the *cnd1-1* mutation. Now that this conclusion had been shown to be false, it seemed essential to delete the *pmt1*$^+$ gene. This analysis is described in chapter 7. Equally important at this stage was the need to establish that the *pmt1*$^+$ gene was expressed and to show that the entire *pmt1*$^+$ gene had been cloned.

As there were only 100bp of DNA 5' to the first ORF of the *pmt1*$^+$ sequence, it was possible that more exons of this gene existed upstream and that the cDNA clone was not a full length copy of the mRNA. One possibility was that the pmt1 protein possessed an N-terminal extension similar to the mouse, human and *Arabidopsis* m^{5}C-MTase enzymes. While the m^{5}C-MTase homologue had been associated with the *cnd1-1* mutation, the complementation of the ts phenotype by the C1 plasmid implied that the 2.4kb *HindIII* m^{5}C-MTase fragment contained the entire gene or at least the functional region of the m^{5}C-MTase. As *pmt1*$^+$ was no longer associated with the mutation, it seemed important to establish whether the entire gene had been identified.

To resolve this uncertainty, Northern blot analysis was performed. A probe made from the *pmt1*$^+$ cDNA clone was hybridized to total yeast RNA and revealed a single transcript of approximately 1.0kb suggesting that the 1073 base pair cDNA clone is likely to be the full length copy of the mRNA (Figure 6.8A). The identification of a strong signal in total RNA suggests that the gene is expressed to a significant level in vegetatively growing cells.

Reverse-transcriptase PCR also confirmed that the *pmt1*$^+$ gene was transcribed and verified the position of the intron (Figure 6.8B). There are two bands in lane 2 of Figure 6.8B which represents the PCR products from RNA treated with reverse transcriptase (RT). The stronger lower band is the size expected for the spliced *pmt1*$^+$ transcript. The upper weaker band is the same size as the band amplified from genomic DNA (lane 2). It is unlikely that the upper band in lane 2 results from the contamination of the RNA sample with genomic DNA, as this band is not present in the PCR reaction that used RNA that had not been treated with RT as a substrate (lane 3). The same sample of RNA was used for both the +RT and -RT reactions. The upper band in lane 2 is most likely derived from unspliced *pmt1*$^+$ mRNA present in total RNA.

6.7 Conclusions

The main conclusion from the work detailed in this chapter is that the *cnd1*$^+$ gene, as defined by the *cnd1-1* mutation, is not the m^{5}C-MTase homologue, thus contradicting the original description of this gene (Bartlett, 1991). The *cnd1*$^+$ gene was correctly identified as *pat1*$^+$ and the *cnd1-1* mutant renamed *pat1-8*. The m^{5}C-MTase homologue was renamed *pmt1*$^+$. In addition, the *scd1*$^+$ gene originally isolated as a multicopy suppressor of

cnd1-1 was found to be identical to the *pac2*$^+$ gene identified by Sugimoto and Yamamoto.

Over-expression of the *pmt1*$^+$ gene does not affect growth rate nor confer any visible phenotype upon wild type *S. pombe* cells. The *pmt1*$^+$ gene is transcribed to a significant level and the size of the mRNA indicates that the cDNA clone corresponds to the entire gene. Having established that the previously identified ORFs represent the complete coding region of *pmt1*$^+$, it was now important to establish whether the *pmt1*$^+$ gene function was essential. The chromosomal deletion of the *pmt1*$^+$ gene is described in the next chapter.

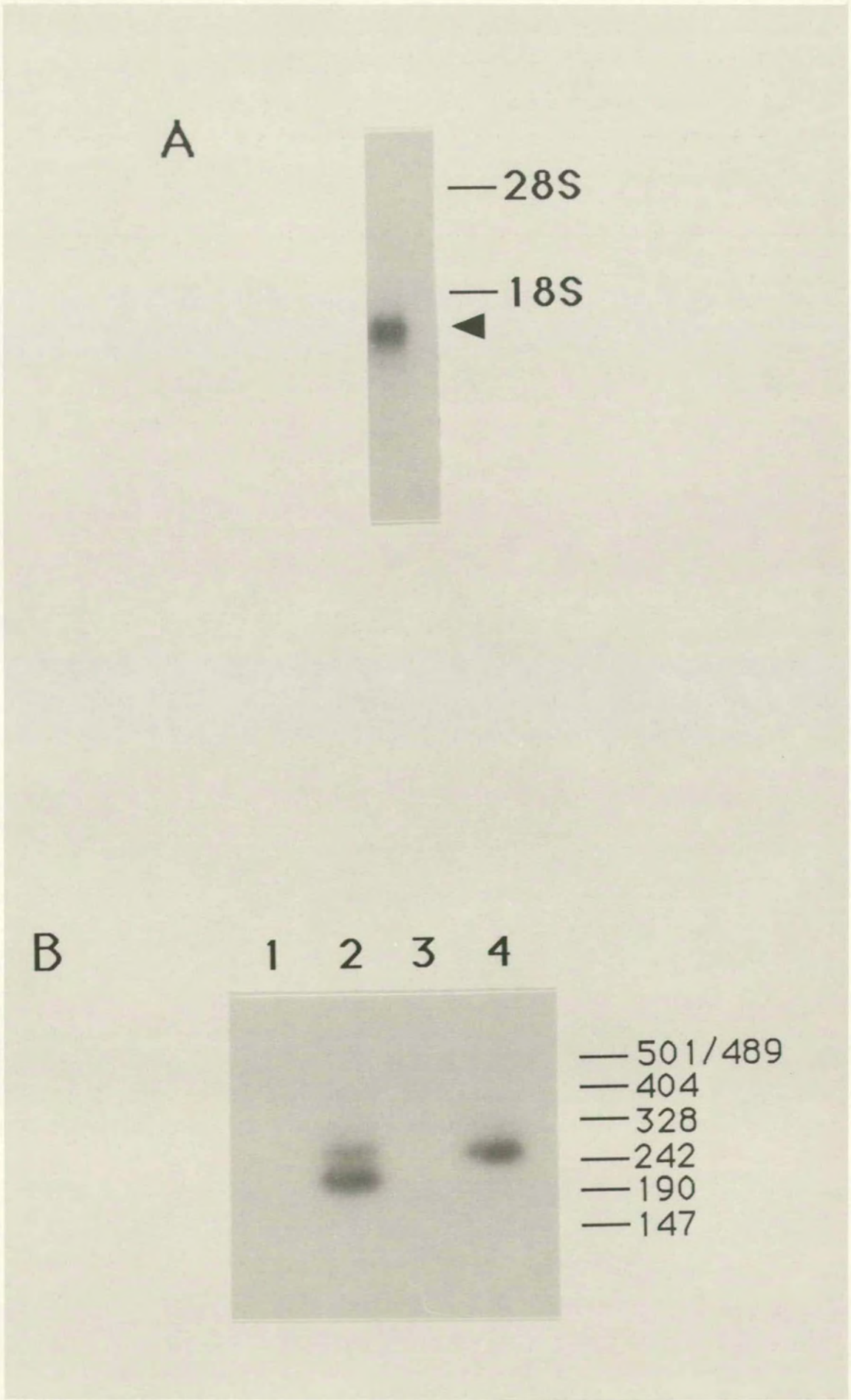
Figure 6.8

A Northern blot analysis to identify the *pmt1*$^+$ transcript

A probe made from the *pmt1*$^+$ cDNA was hybridized to 10μg of total RNA from exponentially growing haploid *S. pombe* cells. The positions of the 28S and the 18S ribosomal RNAs (equivalent to sizes of 3.10 and 1.65 kb respectively, P. Harris, pers. comm.) are indicated. Film was exposed for 2 days at -70ºC.

B Reverse transcriptase PCR analysis of the *pmt1*$^+$ transcript

Total RNA (1μg) from exponentially growing haploid fission yeast cells was incubated with 4μg of oligo dT and reverse transcriptase (RT) as described in section 2.6; a control sample of the same RNA was also set up with no RT. One tenth of these reactions was used as templates in PCR reactions and one tenth of the resulting products run on a 3% agarose gel. PCR reactions were also set up using 50ng of *S. pombe* genomic DNA and a no DNA control. The gel was blotted and probed with the *pmt1*$^+$ cDNA. Film was exposed for 3 days at -70ºC. Molecular weight standards (in base pairs) are shown. The oligonucleotides used in the PCR reaction were designed to span the intron in the *pmt1* sequence and should amplify a 269 base pair fragment from genomic DNA and a 225 base pair fragment from the corresponding reverse-transcribed RNA (details in Appendix A). Lane 1, no DNA. Lane 2, Total RNA +RT. Lane 3, Total RNA -RT. Lane 4, genomic DNA.


A
28S
18S
B
1 2 3 4
501/489
404
328
242
190
147

Chapter 7: Analysis of a *pmt1* null mutation

7.1 Introduction

A powerful way of studying a particular gene function in yeasts is to disrupt or delete the gene at its chromosomal locus and thus create a null mutant. In a disruption experiment, a selectable genetic marker is inserted into a single restriction site within the gene of interest. To delete a gene, the genetic marker is inserted into two restriction sites so that all, or part of the gene is replaced. The one step replacement technique (Rothstein, 1983) is the most commonly used method for creating null mutations. After a suitable marker has been inserted into the gene to be disrupted, a linear fragment is released from the plasmid containing this gene. The fragment should consist of the marker flanked by sufficient length of DNA sequence to permit homologous pairing with both sides of the chromosomal target sequence. Standard techniques are used to transform the linear fragment into yeast cells and select transformants that are stable for the genetic marker.

If the gene in question is essential for cell growth or division, haploid transformants with the required integration event will not be isolated. Hence the replacement is carried out in a diploid strain, so that the majority of transformants are heterozygous for the gene of interest. The heterozygous diploid is then induced to sporulate and the haploid progeny are analyzed. If the disrupted gene is essential, then only two out of the four haploid progeny will be viable. If, on the other hand, the disrupted gene is not essential for growth, then all four haploid spores will give rise to colonies.

In an attempt to study the physiological role of the pmt1 protein, the *pmt1*$^+$ gene was deleted. This chapter describes the construction and

analysis of a strain lacking the *pmt1*$^+$ gene. Prior to such analysis, the chromosomal location of the *pmt1*$^+$ gene was determined to assist in the identification of the required deletion events.

7.2 The chromosomal location of *pmt1*$^+$

i Hybridization of *pmt1*$^+$ to ordered *S. pombe* libraries

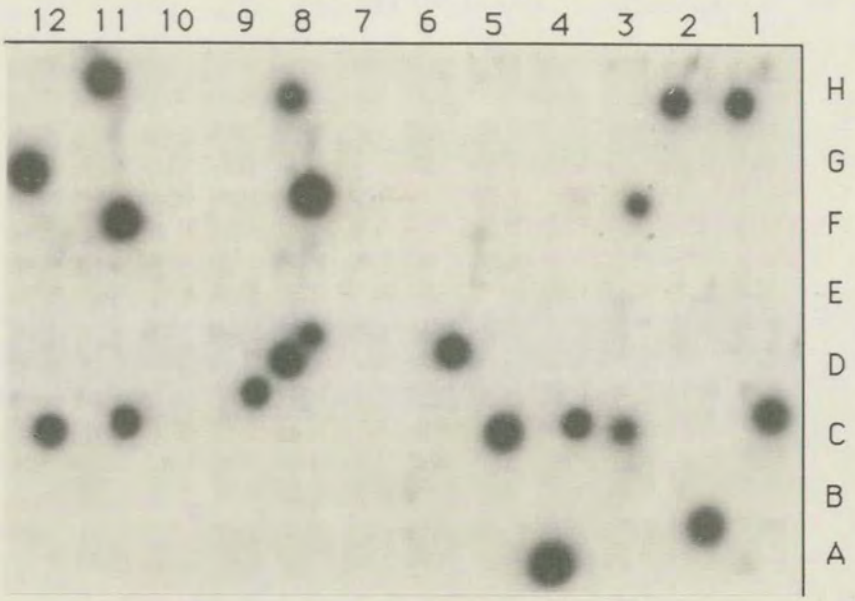
The *pmt1*$^+$ cDNA was hybridized to bacteriophage and cosmid clones covering the entire *S. pombe* genome (Maier *et al.*, 1992; Hoheisel *et al.*, 1993; a kind gift from E. Maier; Figure 7.1). The P1 and cosmid libraries contain 3456 clones each, representing approximately 17 and 8 fold coverage of the *S. pombe* genome respectively. Radio-labelled DNA derived from the cosmid Lorist X was included in the hybridization experiment (a gift from S. Cross; Little and Cross, 1985), such that the ratio of radioactivity between the Lorist X probe and the *pmt1* probe was 1:500. Lorist X shares extensive sequence homology with the cosmid Lawrist 4 which was used for the construction of the cosmid library and limited sequence homology with the bacteriophage P1 vector. The cross hybridization of the Lorist X probe to all the clones gave rise to a weak hybridization signal which facilitated the assignment of hybridization signals to a particular clone within each box.

Strong hybridization signals were obtained to twenty clones on the cosmid filter and to three clones on the P1 filter. The co-ordinates of clones that hybridized to the *pmt1*$^+$ cDNA were kindly assessed by E. Maier (ICRF, London) and revealed, as mentioned in the previous chapter, that the *pmt1*$^+$ gene is located on chromosome II within cosmid 10G12C on the NotI-B fragment. This places the *pmt1*$^+$ gene between the genetic markers *pat1* and *cen2* .

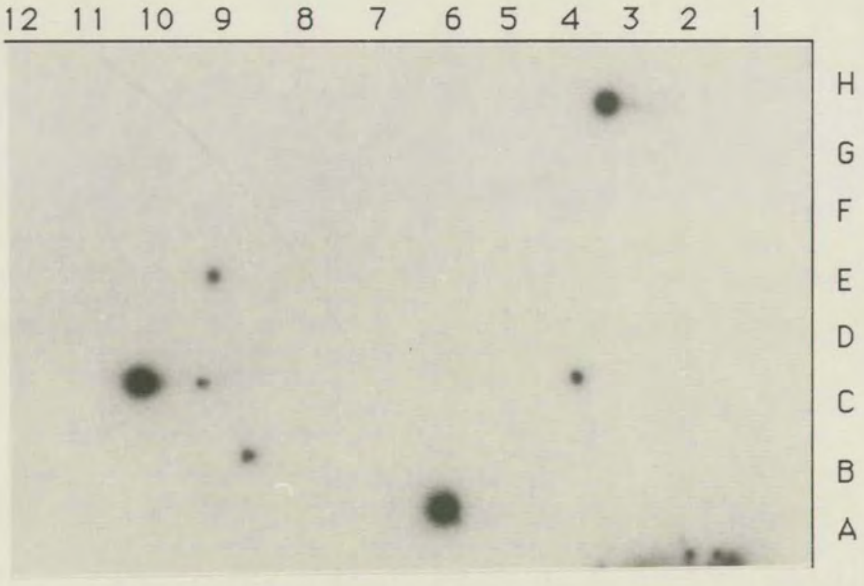
Figure 7.1 Hybridization of *pmt1*$^+$ to *S. pombe* ordered libraries

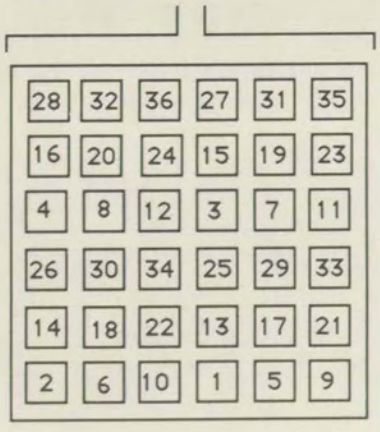
A 1.1kb *HindIII* fragment containing the *pmt1*$^+$ cDNA and the cosmid vector Lorist X were radio-labelled by random hexamer priming and hybridized to high density filters of *S. pombe* ordered libraries in cosmids (A) or in bacteriophage P1 (B). The clones are arranged in 96 boxes (12 x 8), each box contains a 6 x 6 array of clones. The co-ordinates of the clones are defined by a column (A-H), row (1-12) and the location within the box (1-36 as indicated at the bottom of the panel). Film was exposed overnight at room temperature.

A


12 11 10 9 8 7 6 5 4 3 2 1
H
G
F
E
D
C
B
A

B


12 11 10 9 8 7 6 5 4 3 2 1
H
G
F
E
D
C
B
A


28 32 36 27 31 35
16 20 24 15 19 23
4 8 12 3 7 11
26 30 34 25 29 33
14 18 22 13 17 21
2 6 10 1 5 9

7.3 Restriction enzyme mapping of cosmid 10G12C

In order to generate a restriction map of the chromosomal location of *pmt1*$^+$, the cosmid 10G12C (a gift from E. Maier), was digested with a range of restriction enzymes that were already known to cut either the *pmt1*$^+$ or *pat1*$^+$ genomic clones at least once (Figure 7.2). Southern blots of these digests were then hybridized with probes made from the *pmt1*$^+$ cDNA, the *pat1*$^+$ genomic clone and the Lorist X vector (Figure 7.3). Also present on the filters was a *HindIII* digest of the C1 plasmid, which contains the *pmt1*$^+$ and *pat1*$^+$ genomic clones on 2.4kb and 8.5kb fragments respectively.

The insert in cosmid 10G12C was found to be approximately 29.5kb long, of which 22kb was mapped (Figure 7.4), leaving a further 7.5kb corresponding to the ends of the insert DNA positioned adjacent to the Lawrist 4 vector.

During the Southern blot analyses of the *pmt1*$^+$ deletion described later in this chapter (section 7.4), the *pmt1*$^+$ cDNA was hybridized to *S. pombe* genomic DNA cut with many of the enzymes used to generate the restriction map of cosmid 10G12C (Figure 7.4). The bands to which *pmt1*$^+$ hybridized coincided in size with those predicted by the restriction map of cosmid 10G12C (summarized in Figure 7.4), indicating that gross rearrangements of the insert had not taken place. In addition, a clone from a pFL20 library containing this region has subsequently been isolated and the corresponding restriction map found to coincide precisely with that of cosmid 10G12C (M. Baum, pers. comm.).

Figure 7.2 Restriction enzyme analysis of cosmid 10G12C

Cosmid 10G12C was digested with a range of enzymes as indicated for 3 hours at 37°C and then separated by electrophoresis on a 0.8% agarose gel. Each lane contains 0.5µg of cosmid DNA. Molecular weight standards are shown in kilobases.

HindIII
EcoRI
BamHI
SacI
ClaI
Asp718
12.2
11.2
10.2
9.2
8.1
7.1
6.1
5.1
4.1
3.0
2.0
1.6
1.0
0.5

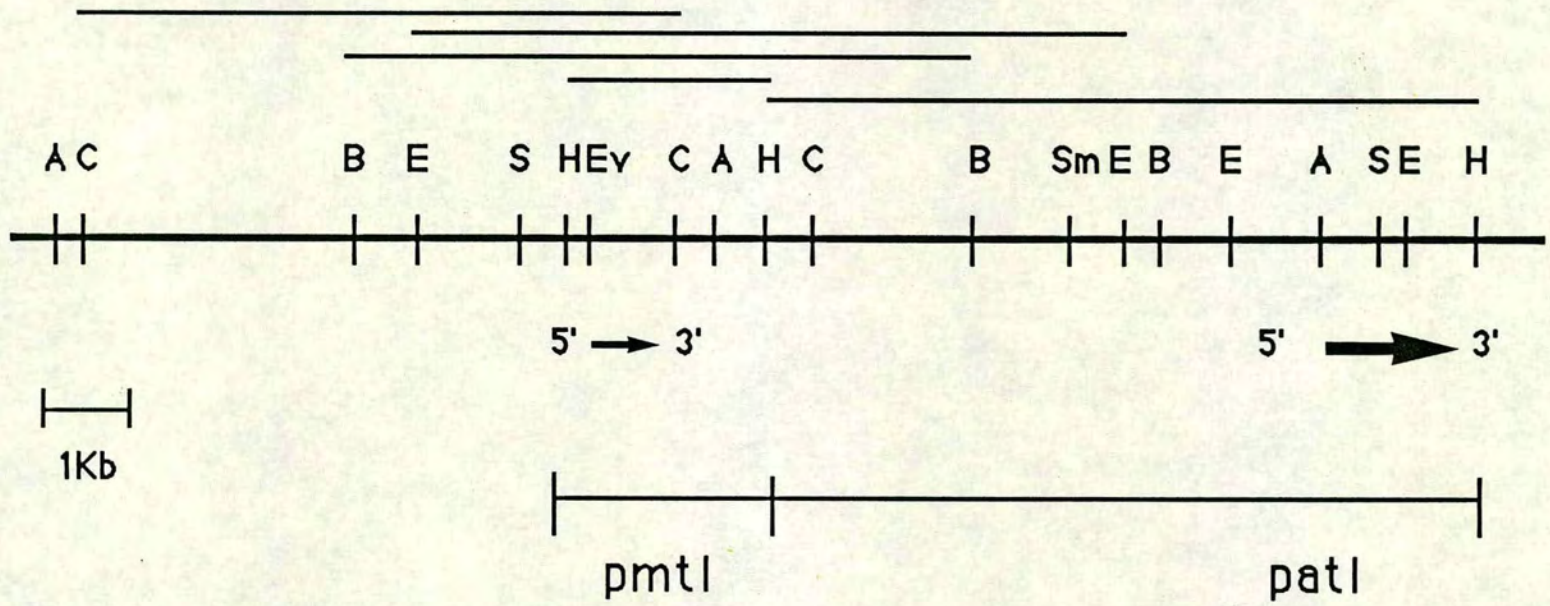
Figure 7.3 Southern blot analysis of cosmid 10G12C

Southern blot of cosmid 10G12C digested with a range of restriction enzymes as indicated. Each lane contains 6ng of cosmid DNA which was separated on a 0.8% agarose gel prior to transfer. Also present on the left hand side of each gel is 20ng of the C1 plasmid digested with *HindIII*. C1 consists of the *pmt1*$^+$ genomic clone and the *pat1*$^+$ genomic clone (which are 2.4kb and 8.5kb *HindIII* fragments respectively) in the 10kb pDB262 vector. Radioactive probes made from Lorist X, the *pmt1*$^+$ cDNA and the 3.5kb *HindIII-BamHI pat1*$^+$ genomic clone were hybridized to the filters. Molecular weight standards are shown in kilobases. Film was exposed overnight at -70ºC.

Lorist X
pmt1
pat1
C1-HindIII
Asp718
ClaI
SacI
BamHI
EcoRI
HindIII
12.2
11.2
10.2
9.2
8.1
7.1
6.1
5.1
4.1
3.0
2.0
1.6
1.0

Figure 7.4 Restriction map of the insert in cosmid 10G12C

A map of approximately 22kb of the 29.5kb insert in 10G12C was generated using the information from restriction enzyme analysis (Figure 7.2) and Southern blot analysis (Figure 7.3). Restriction sites are indicated as follows: A (*Asp718*), C (*ClaI*), B (*BamHI*), E (*EcoRI*), S (*SacI*), H (*HindIII*), Ev (*EcoRV*) and Sm (*SmaI*). The 2.4kb and 8.5kb *HindIII* fragments containing the *pmt1*$^+$ and *pat1*$^+$ genes are shown and the extent and direction of the coding regions illustrated by arrows. The black lines above the map represent the bands identified by Southern blot analysis of genomic DNA using as probes either the *pmt1*$^+$ cDNA or the 3.5kb *BamHI-HindIII* *pat1*$^+$ genomic clone.


A C
B E
S HEv C A H C
B SmE B E A SE H
5' 3'
5' 3'
1Kb
pmtI
patI

7.4 Deletion of the *pmt1*$^+$ gene at its chromosomal locus

i Construction of *pmt1::ura4*$^+$

The pTZ*pmt1::ura4* $^+$ construct used for the *pmt1*$^+$ gene deletion experiment was made by R. Bartlett (University of Oxford) in the following manner. After subcloning the 2.4kb *HindIII* genomic clone of *pmt1*$^+$ into pTZ19R, a 1.2kb *EcoRV-ClaI* fragment containing most of the *pmt1*$^+$ coding region was removed. Using the large fragment of DNA polymerase I, blunt ends were created at *ClaI* site and also on a 1.8kb *HindIII* fragment containing the *ura4*$^+$ selectable marker (Kohli *et al.*, 1977). The *ura4*$^+$ fragment was then ligated into the *pmt1*$^+$ genomic clone (Figure 7.5 A and 7.5 B), with the result that all the *pmt1*$^+$ coding sequence apart from 105bp at the 5' end was removed. This corresponds to the deletion of 294 out of 331 amino acids in the pmt1 protein. In total, there are 200bp of the *pmt1*$^+$ genomic clone at the 5' end and 1020bp at the 3' end of the *pmt1::ura4*$^+$ construct.

A 3kb linear piece of DNA containing the disrupted *pmt1* sequence was generated by PCR using primers located such that the *HindIII* sites at either end of the genomic clone were not included in the product (Appendix B). Consequently, it was possible to use *HindIII* digests to analyze the subsequent deletion events.

Approximately 1µg of the 3kb *pmt::ura4*$^+$ PCR product was transformed by electroporation (section 2.10 v) into a diploid of the genotype *ade6-M210/ade6-M216 ura4-D18/ura4-D18 leu1-32/leu1-32 h*$^+$*/h*$^-$ (details of diploid construction can be found in section 2.10 iii). The two

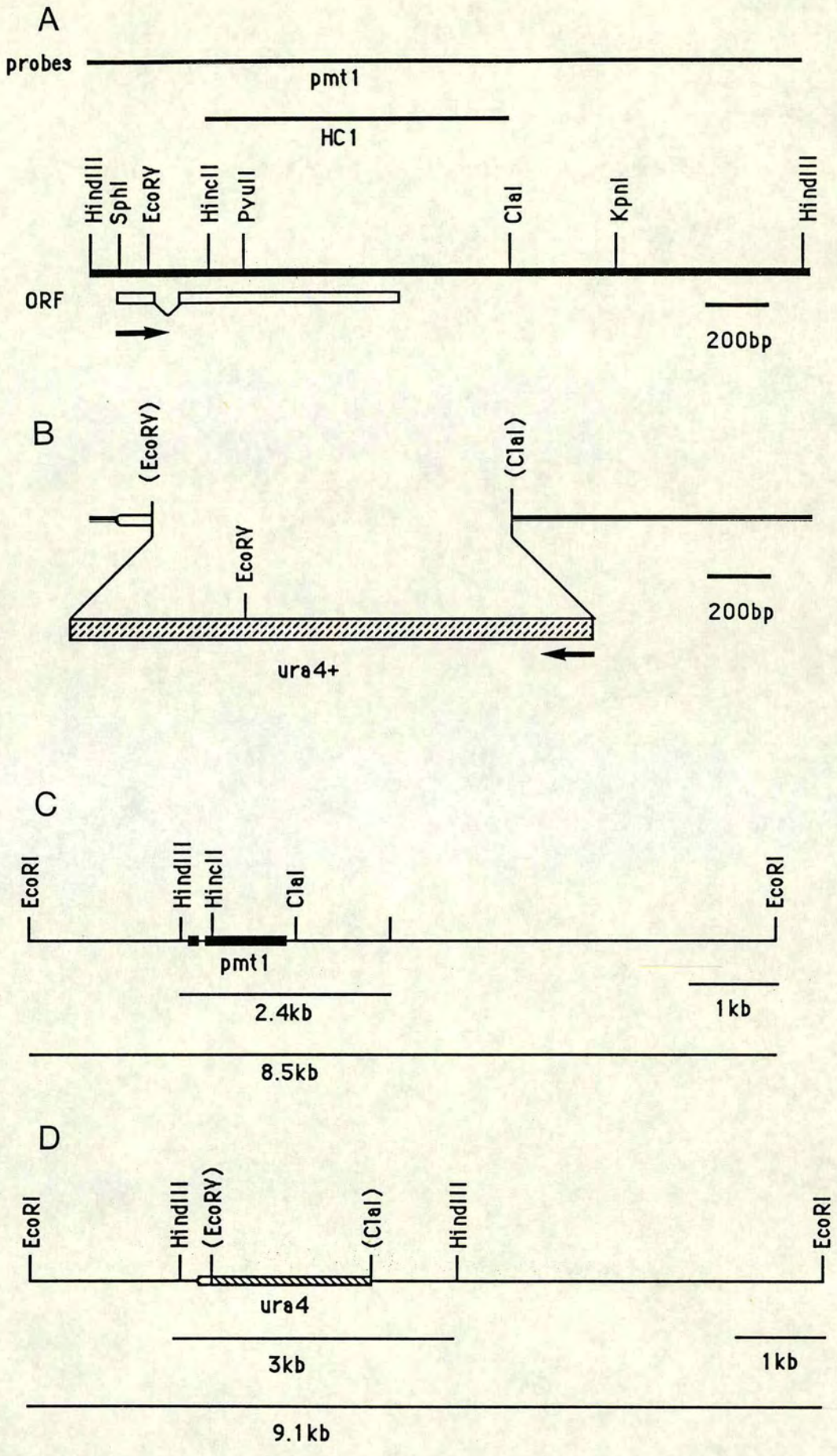
Figure 7.5 Deletion of the *pmt1*$^+$ gene

(A) Restriction map of the *pmt1*$^+$ gene. Shown below the map are the corresponding open reading frames. The probes used for Southern hybridization are shown above the map and are as follows: HC1, the 1kb *HincII-ClaI* fragment; and *pmt1*, the 2.4kb *HindIII* genomic clone. The 1.8kb *ura4* fragment was also used as a probe. The arrow indicates the direction of transcription of the *pmt1*$^+$ gene.

(B) The construction of pTZR19R*pmt1: :ura4*. The *EcoRV-ClaI* fragment of *pmt1* was removed and replaced with the *ura4* gene. The white box represents the part of the *pmt1*$^+$ ORF remaining in *pmt1::ura4*. The arrow indicates the direction of transcription of the *ura4*$^+$ gene.

(C) Restriction map of the 8.5kb *EcoRI* genomic fragment containing the *pmt1*$^+$ gene. The black lines with arrows indicate the size of fragments in which *pmt1*$^+$ is contained (2.4kb *HindIII* and 8.5kb *EcoRI*). The *pmt1*$^+$ ORFs are shown as black boxes.

(D) Restriction map of the 9.1kb *EcoRI* genomic fragment containing the *pmt1::ura4* deletion construct. The black lines with arrows indicate the size of fragments in which the remaining *pmt1* genomic clone is contained (3kb *HindIII*, 9.1kb *EcoRI*). The 1.8kb *ura4* fragment is represented by a striped box. Restriction sites in brackets were lost during the construction of *pmt1::ura4*. The white box represents the part of the *pmt1*$^+$ ORF remaining in *pmt1::ura4*.


A
probes
pmt1
HC1
HindIII
SphI
EcoRV
HincII
PvuII
ClaI
KpnI
HindIII
ORF
200bp
B
(EcoRV)
(ClaI)
EcoRV
200bp
ura4+
C
EcoRI
HindIII
HincII
ClaI
EcoRI
pmt1
2.4kb
1kb
8.5kb
D
EcoRI
HindIII
(EcoRV)
(ClaI)
HindIII
EcoRI
ura4
3kb
1kb
9.1kb

adenine alleles complement one another so that the resulting diploid can grow on media without adenine. Transformants were selected on minimal plates supplemented with leucine. Eight transformants were found to be stable for the *ura4*$^+$ marker.

ii Southern blot analysis to confirm the deletion of *pmt1*$^+$

Southern blot analysis was performed to determine whether the *ura4*$^+$ marker had integrated at the *pmt1* locus in any of the eight stable transformants. The consequences of the expected deletion event are summarized in Figure 7.5 C and 7.5 D. Genomic DNA was made from these strains, digested with *HindIII* and hybridized with a probe made from the *pmt1*$^+$ cDNA. Two of the transformants (Aiii30 and Bi8) had a second *pmt1* band at 3kb; the expected size of the deleted copy of the gene (Figure 7.6). Further blots using DNA digested with *HindIII* and *EcoRI* confirmed that one copy of the *pmt1*$^+$ gene had been deleted in the Aiii30 and Bi8 strains by a single integration event (Figure 7.7). The *pmt1* probe hybridized to a second band in the *EcoRI* digest of approximately 9.1kb giving rise to a doublet with the expected genomic fragment at 8.5kb. The extra *pmt1* band in these blots was also found to hybridize to the *ura4* gene but not to the HC1 probe which represented the deleted portion of *pmt1*. Additional verification of this deletion event was made using blots of *BamHI* and *ClaI* digested DNA (data not shown). Thus the *pmt1*$^+$ gene had been deleted in two diploid strains.

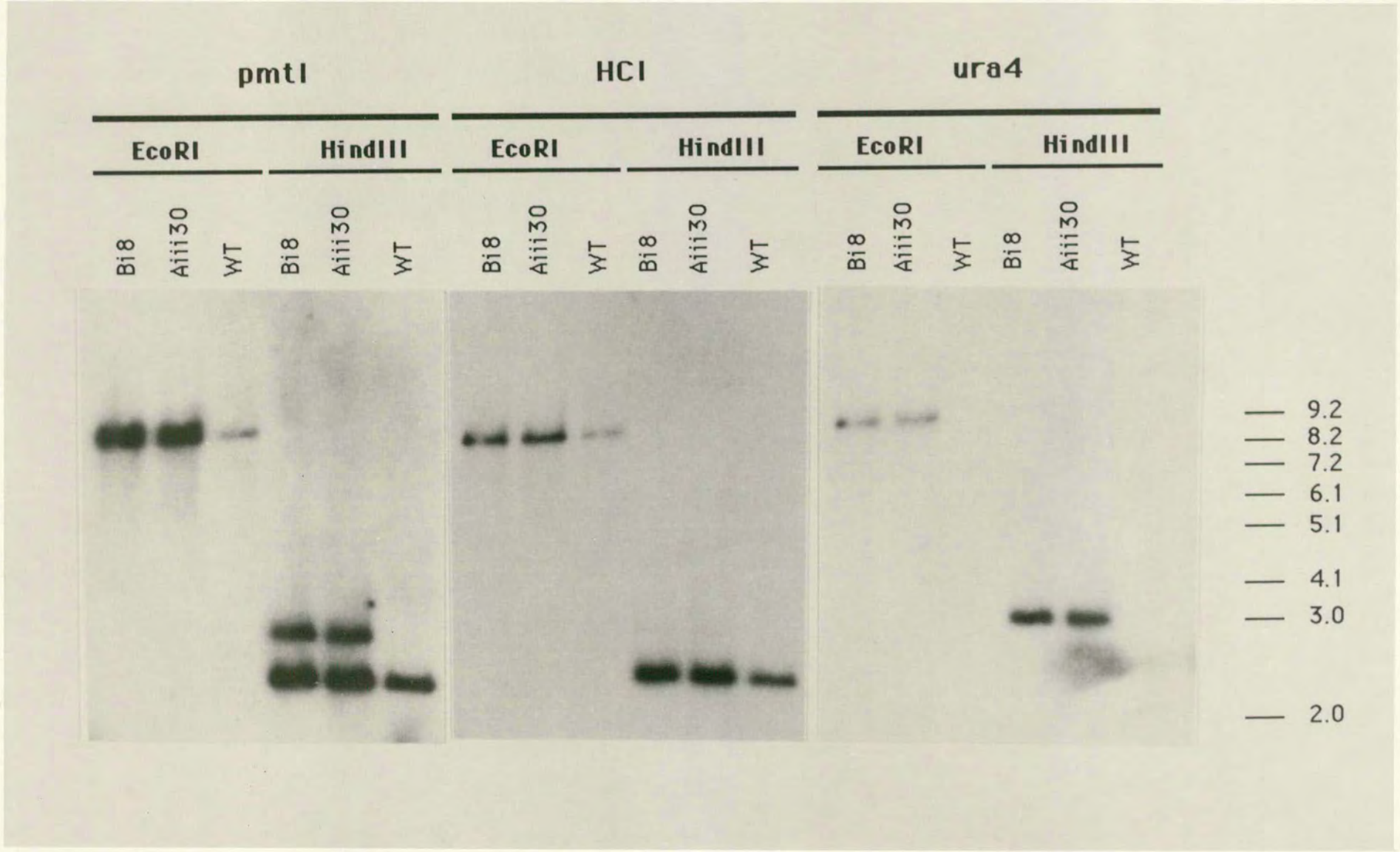
Figure 7.6 Southern blot analysis to assess the integration of *pmt1::ura4*

Southern blot of 0.5μg of genomic DNA isolated from the stable diploid transformants and from a wild type (WT) diploid strain. Prior to transfer to the filter, the DNA was digested with *HindIII* overnight at 37ºC and resolved on a 0.8% agarose gel. The *pmt1*$^+$ cDNA was hybridized to the filter. Film was exposed to the filter overnight at -70ºC. Molecular weight standards (in kilobases) are indicated. Transformants Aiii30 and Bi8 have band patterns consistent with the disruption of one copy of the *pmt1*$^+$ gene (see Figure 7.5).

WT
Bi:i27
Bi:i22
Bi8
Bi6
Av16
Aiv25
Ai:i30
Ai5
10.2
8.2
7.1
6.1
5.1
4.1
3.0
2.0
1.6
1.0

Figure 7.7 Confirmation of the *pmt1*$^+$ gene deletion in transformants Aiii30 and Bi8 by Southern blot analysis

Southern blot of genomic DNA isolated from the stable diploid transformants Aiii30 and Bi8 and also from a wild type diploid strain. Prior to transfer to the filter, 0.5µg of DNA was digested with *HindIII* overnight at 37ºC and resolved on a 0.8% agarose gel. The *pmt1*, HC1 and *ura4* probes were hybridized to the filter as indicated. The same blot was used for each of the three probes. Film was exposed to the filters overnight at -70ºC in each case. Molecular weight standards (in kilobases) are indicated. Transformants Aiii30 and Bi8 have band patterns confirming the deletion of one copy of the *pmt1*$^+$ gene (see Figure 7.5).


pmtI
HCI
ura4
EcoRI
HindIII
EcoRI
HindIII
EcoRI
HindIII
Bi8
Aiii30
WT
Bi8
Aiii30
WT
Bi8
Aiii30
WT
Bi8
Aiii30
WT
Bi8
Aiii30
WT
Bi8
Aiii30
WT
9.2
8.2
7.2
6.1
5.1
4.1
3.0
2.0

iii Tetrad analysis

To analyze the effects of the *pmt1*$^+$ deletion, tetrad analysis was performed on the heterozygous diploids. Initial attempts to sporulate Aiii30 and Bi8, however, were not successful as both strains appeared to be sporulation-defective. The remaining six stable diploids that had undergone non-homologous recombination events were also incapable of sporulation indicating that the sterility was probably not linked to the *pmt1* deletion. The most likely explanation for this observation was that the diploids had become homozygous at the mating type locus. During the propagation of diploids, mitotic recombination can occur at the mating type locus resulting in strains that are either h^+/h^+ or h^-/h^- (Beach and Klar, 1984). These diploids are now stable as both h^+ and h^- gene products are required for sporulation.

While the Southern analysis had been performed, all eight stable diploids had been left at 4ºC. After several days, these strains were re-streaked onto leucine supplemented minimal medium to select for colonies that were still diploid. It seems likely that these events had favoured the selection of diploids that had become homozygous at the mating type locus.

By crossing the diploids Aiii30 and Bi8 to h^+ and h^- haploid strains, it was deduced that the mating types of Aiii30 and Bi8 were h^-/h^- and h^+/h^+ respectively. Diploids that are h^+/h^+ can revert to h^+/h^{90} at a frequency of approximately 10^{-5}. This is due to a rearrangement at the mating type locus (Beach and Klar, 1984). Cells that are h^{90} can switch between the h^+ and h^- states resulting in a colony containing cells of both mating type. This will enable sporulation to occur as this process requires the expression of both the h^+ and h^- genes.

In order to select for an h^+/h^{90} Bi8 strain, cells were plated onto yeast extract medium at a density of 1000 colonies per plate. After three days the colonies were replica plated to malt extract medium to promote meiosis and sporulation and left for a further three days at 28°C. The colonies were then stained with iodine to identify spontaneous conversions of h^+/h^+ to h^+/h^{90}. Colonies that have undergone the conversion will be able to sporulate and will thus be stained black due to accumulation of starch in the walls of spores (Leupold, 1970). Two colonies were stained black by the iodine indicating that they had undergone the transition to h^+/h^{90}. These diploids were chosen from the yeast extract master plate and used for tetrad dissection.

Tetrad analysis was performed in order to analyze the progeny of he Bi8 diploid. A total of 24 tetrads were dissected and in each case all four spores gave rise to viable colonies. Every tetrad displayed 2:2 segregation of ura$^+$ and ura$^-$ haploid cells. Deletion of the *pmt1*$^+$ gene therefore, is not lethal. DNA was made from each of the four progeny from a single tetrad and Southern blot analysis used to confirm the segregation of the *ura4*$^+$ marker with the *pmt1* gene deletion (Figure 7.8). One of the *Δpmt1 ura*$^+$ haploids from this tetrad was used for all further analysis of the *pmt1* deletion.

7.5 Phenotype of the *pmt1* null mutation

i Growth analysis

Further characterization of the strain lacking the *pmt1*$^+$ gene (*Δpmt1*) was carried out to determine whether there were any subtle phenotypes associated with the loss of the *pmt1*$^+$ gene. Cells were grown at a range of

Figure 7.8 Southern blot analysis of a tetrad from the Bi8 diploid

Genomic DNA was isolated from each of the haploid progeny from a single tetrad of the Bi8 diploid. Lanes 1 and 2 contain DNA from the two ura$^+$ progeny and lanes 3 and 4 contain DNA from the ura$^-$ progeny. Each lane contains 0.5µg of DNA that was digested with *HindIII* and separated on a 0.8% agarose gel. The gel was first hybridized to the HC1 probe which represents most of the coding region of *pmt1* and is deleted by the *pmt1::ura4* construct. The blot was then stripped and re-probed with the *ura4* gene. Film was exposed overnight at -70ºC. Molecular weight standards are indicated in kilobases. The band patterns indicate that both ura$^+$ progeny contain the deleted copy and both ura$^-$ progeny contain the wild type copy of the *pmt1*$^+$ gene.

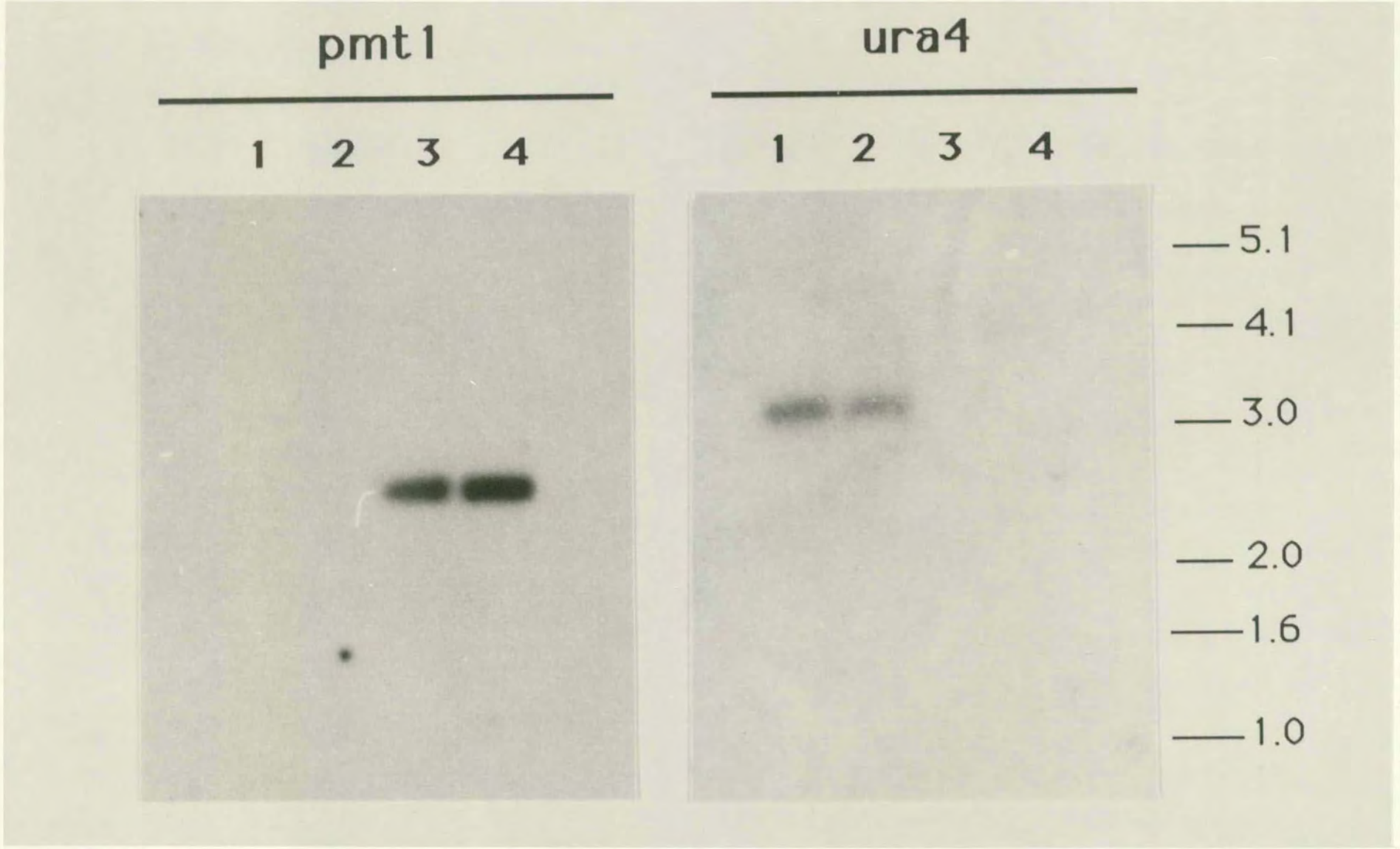
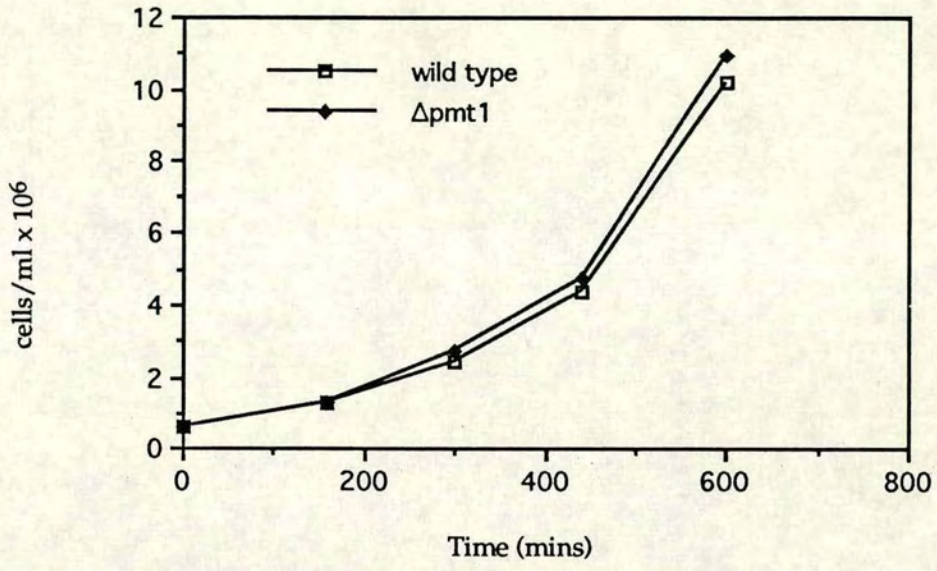

pmt1
1 2 3 4
ura4
1 2 3 4
5.1
4.1
3.0
2.0
1.6
1.0

Figure 7.9 The generation time of *Δpmt1* in complex and minimal media is not significantly different from wild type

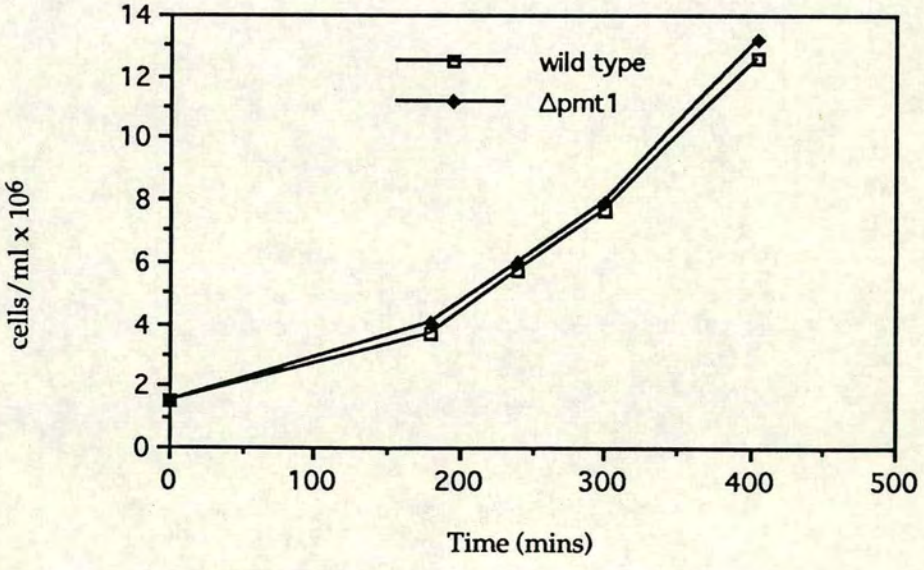
A Wild type and *Δpmt1* cells were grown in yeast extract at 28ºC to a concentration of approximately 1.0×10^7 cells ml^{-1}. The cells were then diluted to approximately 0.6×10^6 cells ml^{-1} in fresh yeast extract and allowed to continue growing at 28ºC. Samples of the culture (100µl) were taken at regular intervals (starting with a time zero sample taken upon dilution), and counted using a Coulter counter. Duplicate samples were taken at each time point and the average value taken to calculate the number of cells ml^{-1}.

B Wild type and *Δpmt1* cells were grown overnight in EMM supplemented with adenine and leucine at 32ºC to an concentration of 1.0×10^7 cells ml^{-1} and then diluted in fresh supplemented EMM to a concentration of 1.4×10^6 cells ml^{-1}. Samples were taken and counted as in A.

A


wild type
Δpmt1
cells/ml x 10⁶
Time (mins)
0
2
4
6
8
10
12
0
200
400
600
800

B


wild type
Δpmt1
cells/ml x 10⁶
Time (mins)
0
2
4
6
8
10
12
14
0
100
200
300
400
500

different temperatures (20, 25, 28, 32 and 36ºC) on complex and minimal media and were examined by light microscopy. No differences between *Δpmt1* and wild type cells were observed (data not shown). The generation time of the *Δpmt1* strain, *pmt1::ura4*$^+$ *ade6-M210 ura4-D18 leu1-32 h*$^+$ was compared with that of *ade6-M210 leu1-32 h*$^+$ at either 28ºC or 32ºC. Both strains had a generation time of 150 minutes in complex medium at 28ºC and 140 minutes in minimal medium at 32ºC (Figure 7.9).

ii Meiosis and mating type switching

Based on the characteristics of DNA methylation and its functions in other organisms, I decided to carry out further investigations with the *Δpmt1* strain. For example, it has been suggested that mating type switching in *S. pombe* involves some sort of epigenetic modification to create asymmetry of the DNA strands at the *mat1* locus (Klar, 1987). Methylation of DNA on one strand could, in theory, provide the required asymmetry and therefore I decided to test whether the *pmt1*$^+$ gene was involved in the regulation of mating type switching.

As described above, fission yeast cells of the opposite mating type (*h*$^+$ and *h*$^-$) will conjugate and sporulate when they are starved of nitrogen. These events can be demonstrated by exposure to iodine which will stain the starch in the spore walls black (Leupold, 1970). Cells that are *h*^{90} can switch between *h*$^+$ and *h*$^-$ states and thus when *h*^{90} colonies are nitrogen-starved, conjugation and sporulation will occur (Egel, 1977). These events will be proportional to the frequency of mating type switching and this can be assessed by subsequent exposure to iodine.

When wild type *h*^{90} and *Δpmt1 h*^{90} colonies were incubated on malt extract medium for three days and then exposed to iodine, the intensity of

staining was the same for both strains indicating that mating type switching was occurring at a similar frequency (data not shown). Deletion of the *pmt1*$^+$ gene therefore, does not markedly affect the efficiency of mating type switching. In addition, this result shows that *pmt1*$^+$ is not required for conjugation, meiosis or sporulation of *S. pombe*.

iii Chromosome stability

It has been reported that a *Neurospora crassa* mutant has reduced levels of methylation and displays abnormal chromosome segregation (Foss *et al.*, 1993; see Introduction). Thus the maintenance of a mini-chromosome by *Δpmt1* cells was analyzed in order to assess the effect of deleting this gene on chromosome stability. A mini-chromosome containing an adenine marker was used in these assays. Cells that lose this chromosome are red due to the accumulation of an adenine precursor since they are *ade*$^-$, giving rise to the appearance of sectored colonies. The frequency at which such colonies arise can thus be used to assess aberrant chromosome segregation events (Hieter *et al.*, 1985).

The mini-chromosome Ch16 is derived from the *S. pombe* chromosome III (Niwa *et al.*, 1986). It is approximately 500kb long and contains the pericentric region of chromosome III including the centromere and three centromere linked markers, *fur1*, *ade6* and *tps16*. The broken ends of duplex DNA are known to be healed by the addition of telomeric sequences and it is thought that this process maintains the length of the mini-chromosome (Haber *et al.*, 1984). Cells containing either the *ade6-M210* or *ade6-M216* alleles require adenine in their growth medium. However, these two alleles can complement one another and cells containing both *ade6-M210* and *ade6-M216* are prototrophic for adenine. A strain that

contains *ade6-M210* on chromosome III and *ade6-M216* on the Ch16 mini-chromosome will grow on an adenine-deficient medium but loss of Ch16 will result in the cells displaying an *ade*⁻ phenotype again.

The mini-chromosome Ch16 (a gift from R. Allshire) of the genotype ***ade6-M216 m23::Leu2*** was crossed into the *Δpmt1* strain to give:
h⁺ pmt1::ura4+ leu1-32 ade6-M210 ura4-D18 Ch16 ade6-M216 m23::LEU2.
As a control, a *pmt1⁺* strain of the following genotype was used:
h⁺ leu1-32 ade6-M210 ura4-D/SE Ch16 ade6-M216 m23::LEU2.

Approximately 20,000 colonies of both the *Δpmt1* and the wild type strain were plated on yeast extract at a density of 1000 colonies per plate. After three days growth at 28ºC the plates were placed at 4ºC for a further three days to develop the intensity of the colour in the sectored colonies. The number of colonies with a red sector covering at least half of the colony were counted. This number divided by the total number of white plus sectored or red colonies gives the number of chromosome loss events per division.

Ten sectored colonies (out of 20000) were obtained with both the *Δpmt1* and the wild type corresponding to a loss rate of 1 in 2000 divisions. This result indicates that the *pmt1⁺* gene has no great effect on the stability of a mini-chromosome and therefore is unlikely to play a major role in controlling the segregation of the three fission yeast chromosomes during cell division.

7.6 Conclusions

This chapter describes the determination of the chromosomal location of the *pmt1⁺* gene and the construction of a *pmt1* null mutation. The *pmt1⁺* gene was found to be located on chromosome II between the centromere

and the *pat1*$^+$ gene. A detailed restriction map of the region containing *pmt1*$^+$ and *pat1*$^+$ was generated by analysis of the cosmid 10G12C.

A *pmt1* null mutant was created by replacing most of the coding sequence of the gene with the *ura4*$^+$ genetic marker. Cells lacking the *pmt1*$^+$ gene were found to be viable and vegetative growth appeared normal. These cells were also capable of conjugation, meiosis, sporulation and mating type switching. In addition, the deletion of this gene was found not to affect to any great extent, the ability of cells to maintain a mini-chromosome.

Chapter 8 Discussion

8.1 Summary of results

The findings described in the previous chapters can be summarized as follows:

1. An analysis of various repeated regions of the *S. pombe* genome was carried out using the isoschizomeric pair of restriction enzymes *HpaII* and *MspI*. The regions assessed included the dg and dh centromeric repeats, telomeres and telomere-associated sequences, the rDNA genes and the LTR of transposable elements. No differences were found between the restriction patterns generated by these enzymes indicating that the internal cytosine of the CCGG sites in these sequences was not methylated. In addition it was not possible to detect methylcytosine within *EcoRII* sites (CCWGG, where W = A/T) by nearest neighbour analysis.

2. Fission yeast DNA was applied to a methyl-CpG binding column but it was not possible to identify a fraction of DNA containing 5-methylcytosine.

3. The multicopy suppressor (*scd1*$^{+}$) of the *cnd1-1* mutation was sequenced and found to encode a small protein with no homology to any of the proteins in current databases.

4. Recombinant histidine-tagged cnd1 (pmt1) protein was purified from *E. coli* and used in a series *in vitro* methyltransferase assays. It was not

possible to detect the transfer of methyl groups by cnd1 to any of the substrates used.

5. Fission yeast cells over-expressing the m^5C-MTase homologue were viable and did not display any visible phenotype.

6. The over-expression of the m^5C-MTase homologue did not rescue the *cnd1-1* mutation. Further analysis revealed that *cnd1*$^+$ is identical to the previously identified gene *pat1*$^+$. The m^5C-MTase homologue was renamed *pmt1*$^+$. The suppressor of *cnd1-1*, *scd1*$^+$, was found to be identical to the previously identified gene, *pac2*$^+$, which is known to suppress mutations in the *pat1* gene.

7. Northern blot analysis revealed that the *pmt1*$^+$ gene is expressed and the size of the corresponding transcript implied that the original cDNA clone was the full length copy of the mRNA. Despite the discovery of significant levels of *pmt1*$^+$ transcript, it was not possible to detect the endogenous pmt1 protein in extracts made from exponentially growing haploid fission yeast cells.

8. The *pmt1*$^+$ gene was found to be located on chromosome II within cosmid 10G12C on the NotI-B fragment. This places *pmt1*$^+$ between the *pat1*$^+$ gene and the *cen2* genetic marker.

9. Haploid cells lacking the *pmt1*$^+$ gene were found to be viable indicating that *pmt1*$^+$ is not an essential gene. Further analysis revealed that the *pmt1*$^+$ gene has no great effect on conjugation, meiosis, mating type switching or chromosome segregation.

The findings summarized above will now be discussed in further detail.

8.2 Methylation analysis of the *S. pombe* genome

Although isoschizomer analysis with *HpaII* and *MspI* has been useful
ethylation in organisms other than yeasts, the detection
is confined to those sites occurring within the recognition
CGG) and thus allows the assessment of only a small
ıcleotides. Indeed, only 4% of cytosines in CpG
cur in CCGG sequences. If a particular sequence is not
the chances of a m^5C residue occurring within a CCGG
remote. Therefore, there could be appreciable amounts
ion in the DNA of a particular organism that will be
e of *HpaII* and *MspI*.

f the limitations of this technique was provided by the
s of the S. *pombe* telomeric region where only one *HpaII*
kb of telomeric repeat and telomere associated sequence
ion, isoschizomer analysis using the *HpaII/MspI* or
s will only detect m^5C that is part of a symmetrical
G or CCWGG) and therefore m^5C residues that are
rical dinucleotide such as CpA, CpT or CpC will go

dvantage of the use of methylation-sensitive restriction
ıethylation status of a particular cytosine is determined
ıtion, namely the lack of cleavage of the DNA. This can
ations regarding the extent of DNA methylation and
rated by the use of *EcoRII* to assay for CNG
3.2). In this experiment, the lack of cleavage by *EcoRII*

gene. *Journal of Molecular Biology* (in press)

was due to reasons other than DNA methylation. The nearest neighbour analysis of *BstNI* digested DNA did not detect a signal corresponding to m^5C. This indicates that less than 0.1% of the internal cytosines in CCWGG sites are methylated. Only approximately 4.8% of CWG sequences will be contained within the *BstNI/EcoRII* recognition site and therefore this analysis does not totally exclude the possibility that CNG methylation exists in the *S. pombe* genome. Furthermore, this analysis corroborated the findings of others (Pein *et al.*, 1991; Kruger *et al.*, 1988) in observing that *EcoRII* is inefficient at cleaving genomic DNA.

A column containing the methyl-CpG binding domain of the rat protein MeCP-2 has been used to separate DNA from various organisms into methylated and non-methylated fractions (Cross *et al.*, 1994; S. Cross and J. Charlton, pers. comm). When *S. pombe* DNA was applied to the MBD column, it was not possible to detect a methylated fraction (section 3.4) implying that there are no repeated regions of the *S. pombe* genome that are subjected to cytosine methylation at CpG dinucleotides. However, this method is not sensitive enough to detect single copy sequences. One way to extend this analysis would be to probe the fractions predicted to contain methylated DNA with particular single copy sequences. If any genes were found in these fractions, their methylation status could be confirmed by other methods, such as isoschizomer analysis or genomic sequencing.

Despite the failure of attempts to detect 5-methylcytosine in fission yeast, it is still possible that this modification exists. Even the most sensitive techniques require that 0.075% of the total cytosine is modified (Crain and McCloskey, 1983). Consequently, there could be at least 2100 undetected m^5C residues in the fission yeast genome (assuming a 40% GC content and a genome size of 14Mb).

Genomic sequencing of bisulphite treated DNA has many advantages over other forms of analysis for detecting methylation (Frommer *et al.*, 1992; section 1.2). It results in a positive signal for m^5C, does not require cytosine to be part of a symmetrical sequence and it allows the identification of hemi-methylated sites not normally detected with restriction enzymes. Although the bisulphite method is time-consuming and not without technical difficulties, this would be the most sensitive way to determine whether certain regions of the *S. pombe* genome were methylated. Candidate regions for such an investigation are considered below (section 8.5).

Perhaps the easiest way to identify rare sites of DNA methylation in *S. pombe* would be to digest genomic DNA with a methylation-dependent restriction enzyme and then analyze the digestion products using pulse field gel electrophoresis (PFGE). This latter technique is able to resolve the three *S. pombe* chromosomes (Smith *et al.*, 1987) and even if there were only a few methylated sites in the genome, a different gel pattern would easily be seen. A m^5C dependent restriction enzyme has been identified (Ross *et al.*, 1989; Dila *et al.*, 1990). The *mcrBC* locus of *E. coli K-12* encodes a restriction enzyme that mediates restriction of cytosine modified DNA including sequences containing 5-hydroxymethylcytosine and N^4-methylcytosine as well as m^5C. Two proteins are encoded by the locus; a restriction component McrB whose specificity is modified by the second protein McrC (Dila *et al.*, 1990). Both proteins have been purified and their specificity of restriction has been analyzed *in vitro* leading to the determination of the consensus motif $R^mC(N40\text{-}80)R^mC$. Cleavage occurs at multiple sites on both strands between the modified cytosines (Sutherland *et al.*, 1992).

8.3 Failure to detect DNA methyltransferase activity

The recombinant histidine-tagged pmt1 (formerly cnd1) protein was over-expressed in *E. coli*. Moderate levels of expression were obtained and the resulting cell extracts used to purify the pmt1 protein using nickel ion affinity chromatography. A series of *in vitro* methyltransferase assays were performed, but it was not possible to detect transfer of methyl groups by pmt1 to any of the substrates used. The lack of detectable activity could indicate that the protein is not a functional m^5C-MTase, a hypothesis consistent with the lack of detectable methylation in this organism. On the other hand, if the pmt1 protein is an active methyltransferase, there could be several reasons why it was not possible to demonstrate this.

Pmt1 may require certain modifications *in vivo*, such as glycosylation, phosphorylation or even association with other subunits or cofactors. Although the recombinant pmt1 was incubated with *S. pombe* protein extracts, this may not have been sufficient for any such modifications to occur. Another possible explanation for the apparent inactivity could be that the recognition sequence of pmt1 is more than seven base pairs long. The substrates that were used may not have contained enough copies of this sequence to give a detectable signal when methylated with ^3H-AdoMet.

The recombinant protein contains a six-histidine tag, but it is unlikely that this interferes with the protein structure as other m^5C-MTases are known to be active when fused to GST subunits at their N-termini (Taylor *et al.*, 1993). Moreover, the length of N-terminal sequence in m^5C-MTases that precedes motif I is highly heterogeneous and X-ray crystallography studies have shown that the N-terminal tail is likely to be located on the outside of the protein and not buried inside with the conserved motifs (Cheng *et al.*, 1993).

8.4 The *S. pombe* m^{5}C-MTase homologue is not the *cnd1*+ gene

A number of fundamental errors led to the incorrect assignment of the *S. pombe* m^{5}C-MTase gene as *cnd1*+ (Bartlett, 1991). If the *HindIII* digest of the C1 plasmid (Figure 6.3A) had not been sufficiently resolved by gel electrophoresis, the 8.5kb *pat1*+ insert would have run as a doublet with the 10kb vector band. As the plasmid was expected to contain only one *HindIII* insert, it was presumed that the 2.4kb band contained the *cnd1*+ gene.

Transformation of the C1 plasmid into the *cnd1-1* mutant resulted in some colonies displaying stable rescue. Analysis of these colonies indicated that the plasmid had integrated by homologous recombination at the *cnd1-1* locus and it was concluded that C1 contained the *cnd1*+ gene (Bartlett, 1991). Had the corresponding cDNA been cloned by complementation rather than by hybridization to the genomic fragment, the correct identification of *cnd1*+ would have been made. In addition, had the C1 plasmid been digested with another restriction enzyme in the original analysis, the extra insert would have been noticed.

A similar set of incorrect conclusions was made with the cloning of the *MAD2* gene from *S. cerevisiae* (Li and Murray, 1991). The *mad2-1* ts mutation is defective in the feedback control which ensures that the exit from mitosis is dependent on the completion of spindle assembly. A plasmid from a genomic library was isolated that was able to complement the mutation and integrate at the *mad2* locus. Contained within the genomic clone was an ORF that showed homology to the α-subunit of prenyltransferases. To further investigate the role of the MAD2 protein, a new ts mutant, *mad2-2* was isolated (Li *et al.*, 1993) and was found to have a strikingly different phenotype to *mad2-1* (Li *et al.*, 1993).

Subsequent analysis of these mutants led to the realization that the prenyltransferase was adjacent to the bona fide *MAD2* gene (Li *et al.*, 1994). With both *cnd1* and *MAD2*, the failure to demonstrate complementation of the ts mutations by a cDNA clone led to major misinterpretations of results.

After discovering that the *cnd1*$^+$ gene was *pat1*$^+$, it transpired that 17 out of the 20 mutations identified in the same screen used to isolate *cnd1-1* were also in the *pat1* gene. These alleles were thought to correspond to the more severe and previously identified mutation *pat1-114*. Unfortunately, the *cnd1-1* mutant was not identified as being a *pat1* mutant at this time as its phenotype appeared different from those displayed by *pat1-114* (P. Nurse, pers. comm).

The original interpretation of the effects of the *cnd1-1* mutation was that cells became blocked in nuclear division leading to an increase in DNA content. However with hindsight, it seems more likely, given the known functions of *pat1*$^+$, that the phenotype is a result of cells attempting to undergo meiosis whilst still in the haploid state. There are ts mutations that are less severe than *pat1-114* including *pat1-3* and *pat1-4* (Nurse, 1985), and it is possible that *pat1-8* is identical to one of these alleles.

The *scd1*$^+$ gene was sequenced in an attempt to gain information about the *cnd1*$^+$ gene, but was eventually found to be identical to the previously identified gene, *pac2*$^+$. This gene was sequenced in 1989 (M. Yamamoto, pers. comm) and had its sequence been in the database, the identity of the *cnd1*$^+$ gene might have been resolved earlier on in the course of this project.

The *pat1*$^+$ protein kinase (also known as *ran1*$^+$) is a critical negative regulator of meiosis and its inactivation is both necessary and sufficient to cause entry into the meiotic cell cycle (Iino and Yamamoto, 1985; Nurse, 1985). In fission yeast, meiosis is initiated by the *mei3*$^+$ gene, the

transcription of which is induced by the products of the mating type genes from the *mat1-M* and *mat1-P* loci (McLeod *et al.*, 1987). The mei3 protein is thought to initiate meiosis by binding non-covalently to the pat1 protein kinase, thereby effecting its inactivation (McLeod and Beach, 1988). The product of the *mei3*$^+$ gene is a small protein of 148 amino acids and shows no homology to any sequence in the protein databases. However, analysis of the amino acid content reveals that 15% of residues are serine and 15% are threonine and thus mei3 is similar in composition to the predicted pac2 protein sequence. One possible mechanism by which *pac2*$^+$ is able to rescue the *pat1* ts mutations is that it competes with mei3 for binding to pat1, but whereas mei3 inactivates the protein kinase, pac2 binding permits activity and thereby ensures that the cells remain in the mitotic cell cycle. Alternatively, pac2 could be a substrate for pat1. Phosphorylated pac2 protein could then be a component in the pathway by which active pat1 mediates commitment to the mitotic cell cycle.

8.5 Analysis of the *pmt1*$^+$ null mutant

Haploid cells lacking the *pmt1*$^+$ gene were found to be viable (section 7.4 iii) and appeared indistinguishable from wild type cells in vegetative growth. Furthermore, *Δpmt1* cells were found to be capable of conjugation, meiosis, sporulation, mating type switching and the correct segregation of a mini-chromosome.

In addition to these experiments, the sensitivity of the *Δpmt1* strain to 5-azacytidine has been assessed. These experiments were carried out by R. MacDonald and C. Price (University of Sheffield). The analogue 5-azacytidine (5-azaC), has been used by many investigators to study the effects of DNA methylation by inducing hypomethylation. This drug is an

analogue of cytosine and upon application to cells, can become incorporated into DNA. Its mechanism of action has not been clearly elucidated but it is known to be able to bind irreversibly to m^5C-MTases (Christman *et al.*, 1985; Santi *et al.*, 1984) thus preventing further catalysis by these enzymes. This leads to a reduction in DNA methylation.

Preliminary experiments have indicated that certain doses of 5-azaC can lead to an arrest at the G2/M boundary in *S. pombe* cells. Although the mode of action is unclear, the delay appears to be mediated through the DNA damage checkpoint control (C. Price, pers. comm). These effects could be explained by the irreversible binding of a m^5C-MTase to the 5-azaC. Such a permanent MTase-DNA complex might well be expected to hinder mitosis and would be registered by the cell as a form of DNA damage.

If the *pmt1*$^+$ gene does encode a functional m^5C-MTase, then one might expect cells lacking the pmt1 protein to be less sensitive to 5-azacytidine. The response of the *Δpmt1* strain to 5-azaC was analyzed but no noticeable difference between the response curves of the *Δpmt1* and wild type strains were found indicating that it is unlikely that the pmt1 protein is causing the cell cycle delay seen in response to 5-azaC treatment (C. Price, pers. comm). There have also been reports suggesting that 5-azaC can also form stable complexes with nuclear proteins other than m^5C-MTases (Christman *et al.*, 1985). This mechanism could also account for the cell cycle arrest seen upon treatment of *S. pombe* cells with 5-azaC.

The *Δpmt1* strain could be used in further experiments to try and elucidate the function of this gene. For example, a possible role of *pmt1*$^+$, could be to methylate DNA to distinguish existing DNA strands from newly replicated ones. This could enable a mismatch repair system to operate in a manner analogous to the role of adenine methylation in bacteria (Modrich, 1991). Such a mechanism does not require dense methylation consistent with

the lack of detectable m^5C in *S. pombe*. In order to test this theory, a G-T mismatch could be incorporated into a plasmid such that correction to G-C or A-T would create different restriction sites. A comparison of the nature of correction events between wild type and *Δpmt1* strains could then be made to determine whether *pmt1* is involved in such a mechanism.

DNA methylation in eukaryotes is mainly associated with the silencing of gene expression. The mating type locus in fission yeast is known to contain regions of transcriptionally inert DNA, namely the *mat2* and *mat3* genes. These genes act as stores of information that are periodically transposed into the active *mat1* locus during mating type switching. It is not known how the *mat2* and *mat3* genes are repressed although it is thought that unusual chromatin structure may contribute to the silencing mechanism as meiotic interchromosomal recombination in *mat2*, *mat3* and the intervening 15kb region does not occur at all (Egel, 1984). In addition, other genes that are introduced nearby are also silenced (Thon and Klar, 1992; Thon *et al.*, 1994).

One way to test whether $pmt1^+$ is involved in the regulation of silencing at *mat2* and *mat3* would be to check for expression of *mat3* in an h^{+N} strain. This strain contains both *mat2* and *mat3* information but will only express *mat2* at the active loci due to a rearrangement of the entire *mat* locus, Beach and Klar, 1984). The inhibition of recombination and transcription in this region has been shown to depend on several trans-acting loci, *clr1*, *rik1* and *swi6* (Thon and Klar, 1992; Lorentz *et al.*, 1992) and also *clr2*, *clr3* and *clr4* (Thon *et al.*, 1994; Ekwall and Ruusala, 1994). In addition, four cis-acting elements which repress *mat2* in a plasmid have been identified (Ekwall *et al.*, 1991). Deletion of two of these elements from the chromosome has minimal effect. However in combination with mutations in some of the trans-acting factors described above, deletion of the same cis-acting elements was found

to greatly enhance *mat2* expression (Thon *et al.*, 1994). These findings imply that there is a redundancy in the silencing mechanism. Double mutants could be created with some of these mutations and the *Δpmt1* strain to determine whether *pmt1*$^+$ contributes to the silencing mechanism.

Although the deletion of the *pmt1*$^+$ gene was found to have no great effect on mating type switching, it may still be involved in the regulation of this process, perhaps in combination with other genes. Deletion of the *pmt1*$^+$ gene could alter the effects of mutations in genes involved in the regulation of switching. For example, mutations in the *swi* genes, *swi1*, *swi3* and *swi7* show a reduced level of double strand break at *mat1* (Egel et al., 1984), whereas *swi2* and *swi6* mutants possess normal levels of the break yet they still fail to switch efficiently (Gutz and Schmidt, 1985). The *swi6*$^+$ gene has been shown to encode a homologue of chromatin associated proteins from *Drosophila* and mammals and therefore is likely to be involved in the control of chromatin structure in this region (Lorentz *et al.*, 1994). All *swi* mutants still show some mating type switching, once again suggesting a redundancy in the system. In order to determine whether *pmt1*$^+$ plays some sort of role in the regulation of mating type switching, the extent of switching could be assessed in double mutants containing the *pmt1* deletion and the *swi* mutant alleles.

Other areas of the *S. pombe* genome known to contain unusual chromatin structure are the centromeres and telomeres. Insertion of the adenine or uracil markers into these regions results in the repression of expression and the spreading of the nearby chromatin structure into those genes (Allshire *et al.*, 1994; Nimmo *et al.*, 1994), resembling the phenomenon of position effect variegation seen in *Drosophila* and *S. cerevisiae*. An interesting experiment would be to compare the reduction of marker expression at the centromeres and telomeres between a wild type and the

Δpmt1 strain. Due to the presence of unusual chromatin structure and the consequent silencing of gene expression, regions of the *S. pombe* centromeres, telomeres and mating type loci would be good candidates for a methylation analysis by genomic sequencing.

8.6 The significance of the serine-cysteine peptide at the catalytic site

Although pmt1 conforms well to the consensus for m^5C-MTase enzymes, one remarkable divergence from the consensus stands out. The pmt1 protein sequence has a serine adjacent to the cysteine in the catalytic site, whereas a proline is found in all other m^5C-MTases. The proline-cysteine dipeptide appears to be conserved, not only in m^5C-MTases but also in other enzymes that effect methyl group transfer by the same catalytic mechanism, such as thymidylate synthase. During catalysis, the cysteine is known to provide the thiol group as a nucleophile. The role of the proline however, is unclear. From the X-ray crystallography studies of the M.HhaI m^5C-MTase (Cheng *et al.*, 1993; Klimasauskas *et al.*, 1994), it would appear that the amide carbonyl preceding the proline hydrogen bonds to the exocyclic amine of the target cytosine. The carbonyl group of the proline itself seems to form a hydrogen bond with the side chain of the asparagine in the conserved ENV sequence in motif VI. The proline is situated around the hydrophobic pocket where the cofactor AdoMet binds, and its hydrophobic side chain may contribute to this environment.

Nevertheless, from these interactions, it is hard to see why this residue is so conserved. One possibility is that the proline may be restricting the possible conformations of the adjacent cysteine and thus many

alternative amino acids will not fulfil this role. To date, there have been no mutagenesis studies to address this question.

One possible explanation for the presence of the serine instead of proline in pmt1 is that the serine has become inserted in between proline-79 and cysteine-81 thereby inactivating what was previously a functional m^5C-MTase. However, as the preceding residues in this motif have not been completely conserved, it is not possible to judge the likelihood of this. M.HgaI-2 and the mouse enzyme have the sequence PPC in this region and therefore pmt1 could have undergone a change in sequence from PPC to PSC. This change would have required at least two point mutations to derive the present serine codon AGT, from the codons for proline.

8.7 The possible biological significance of *pmt1+*

As the *pmt1+* gene is non-essential and the purified recombinant pmt1 protein displays no detectable m^5C-MTase activity, it could be concluded that the *pmt1+* gene function is not required in fission yeast at all, and that it is merely an evolutionary relic; a hypothesis that is consistent with the lack of detectable methylation in this organism (Antequera *et al.*, 1984; Chapter 3). This theory is also supported by the presence of the divergent serine-cysteine dipeptide in the catalytic site.

On the other hand, the *pmt1+* gene is transcribed at a significant level in vegetatively growing cells (section 6.6) and the predicted protein still contains the catalytic cysteine and all ten conserved motifs. These observations argue that *pmt1+* may encode a functional m^5C-MTase. However, it has not been possible to detect the pmt1 protein in total cell extracts using polyclonal antisera generated against the purified pmt1

protein. This may have been due to the quality of the antibodies or to the presence of only small amounts of the protein.

Alternatively, the *pmt1*$^+$ transcript may only be translated at certain stages of the life cycle, for example during meiosis or sporulation. If *pmt1*$^+$ does encode a functional m^5C-MTase, one would expect to find the protein localized in the nucleus. The pmt1 protein sequence does not possess a strong nuclear localization signal, but its position in the cell could be determined by association with other factors. One way to resolve this issue would be to tag the pmt1 protein with an epitope and follow its cellular distribution by indirect immunofluorescence.

The function of the *pmt1*$^+$ gene may be required only under certain conditions, hence its dispensability for vegetative growth. Fission yeast cells that are grown in a laboratory are not subject to the same environmental stresses that they might encounter in their natural environment. Pmt1 could be part of a yeast restriction-modification system that has been rendered redundant by cultivation for many generations under laboratory conditions. An interesting experiment would be to obtain *S. pombe* that is growing in the wild and analyze the methylation content of its DNA.

Alternatively, as transposons have been identified in *S. pombe* (Levin et al., 1990), methylation could control the spread of these elements through the yeast genome (even though methylation was not detected at CG dinucleotides within CCGG sequences in these elements; section 3.3). A process that detects and inactivates duplicating sequences has been detected in other fungi (RIP, MIP; see Introduction). However, it is unlikely that a process exactly analogous to RIP and MIP exists in *S. pombe* as duplicated sequences that are transformed into the genome (such as nutritional markers) do not become inactivated.

One further possibility is that *pmt1*$^+$ is a pseudo-gene and that another functional methyltransferase exists. Low stringency southern hybridizations using *pmt1*$^+$ as a probe have not provided any evidence in favour of this suggestion (data not shown). An alternative way to search for related genes, would be to perform degenerate PCR using primers complimentary to two of the conserved regions. The *Arabidopsis* m^{5}C-MTase gene was cloned in this way (Finnegan and Dennis, 1993) and a potential m^{5}C-MTase gene family has subsequently been identified (Scheidt *et al.*, 1994).

Even if DNA methylation does exist in *S. pombe*, there is clearly only a very small proportion of cytosine residues that are modified. It is clear though, that the presence or absence of small numbers of methyl groups can have a disproportionately large biological effect. For example, two methylated cytosines on opposite strands of a short palindromic sequence can prevent DNA from restriction by certain enzymes. In addition, the mammalian protein MeCP-2 selectively binds to DNA containing just one symmetrically methylated CpG pair (Lewis *et al.*, 1992) thereby potentially preventing alternative protein-DNA interactions at this site.

8.8 Could pmt1 be an RNA methyltransferase?

As described above, the major difference between pmt1 and all other m^{5}C-MTases cloned to date is the presence of a serine-cysteine dipeptide in the active site. Intriguingly, a serine-cysteine dipeptide has been seen at the active site in a methyltransferase that modifies uracil residues in tRNA. Like the m^{5}C-MTases and thymidylate synthase, tRNA uracil methyltransferase catalyzes the transfer of methyl groups onto the 5-position of a pyrimidine (Santi and Hardy, 1987). This suggests that a serine next to a cysteine is not

necessarily inhibitory for this type of reaction and it raises the possibility that the natural substrate of pmt1 may be an RNA. The methylation of cytosine at the 5-position has been found in ribosomal RNA in organisms ranging from bacteria to humans (Noller, 1984; Maden, 1988) but the cloning of the enzymes responsible for this modification has not been reported. Attempts to methylate *S. pombe* RNA (Chapter 5) were not successful but it could be that the available sites in rRNA or even in mRNA were fully modified. At the time that these experiments were carried out it was believed that the *cnd1-1* mutation was associated with the m^5C-MTase homologue, and therefore RNA was made from the mutant and also used as a substrate in these assays. With hindsight, this RNA was also effectively wild type. A more valid experiment would be to isolate total RNA from the *Δpmt1* strain and use this as a substrate for *in vitro* methyltransferase assays.

Further *in vitro* methyltransferase assays have been carried out in conjunction with site-directed mutagenesis studies of the pmt1 protein (D. Hornby, University of Sheffield). The serine residue adjacent to the cysteine in region IV was replaced by a proline to give the PC consensus dipeptide. The mutated pmt1 protein was then apparently found to possess DNA methyltransferase activity using the tritium transfer assay. While this experiment reinforces the view that pmt1 has probably evolved from an active m^5C-MTase, the specificity and function of pmt1 remains unresolved. Certainly, this finding greatly favours the view that *pmt1*$^+$ encodes a redundant m^5C-MTase. Interestingly, a single amino acid change (Lys188 to Gln188) also restores m^5C-MTase activity to a *chlorella* virus pseudogene (Zhang *et al.*, 1992). However, the substitution of the proline by serine in pmt1 could merely be a way of reducing activity. Alternatively, the serine could be responsible for conferring upon pmt1, the ability to methylate RNA. One way to determine whether pmt1 is a DNA methyltransferase

with extremely low activity would be to use the active pmt1 m^5C-MTase with the serine to proline mutation, to determine its DNA recognition sequence. Once this information is known, further assays could be carried out whereby the wild type protein is incubated with oligonucleotides containing the relevant sequence thereby greatly increasing the resolution of the methyl transfer assay. These experiments are currently being undertaken (D. Hornby, pers. comm.).

References

Adams, R.L.P and Burdon,R.H. (1985) Molecular biology of DNA methylation. Springer Verlag. New York.

Allshire,R., Javerzat,J.P., Redhead,N.J. and Cranston.G. (1994) Position effect variegation at fission yeast centromeres. *Cell*, **70**, 157-169.

Antequera,F., Tamame,M., Villanueva,J.R. and Santos,T. (1984) DNA methylation in the fungi. *J. Biol. Chem.*, **259**, 8033-8036.

Antequera,F., Tamame,M., Villanueva,J.R. and Santos,T. (1985) Developmental modulation of DNA methylation in the fungus *Phycomyces blakesleeanus*. *Nucleic Acids Res.*, **13**, 6545-6588.

Antequera,F. and Bird,A. (1988) Unmethylated CpG islands associated with genes in higher plant DNA. *EMBO J.*, **7**, 2295-2299.

Antequera,F., Boyes,J. and Bird,A. (1990) High levels of *de novo* methylation and altered chromatin structure at CpG islands in cell lines. *Cell*, **62**, 503-514.

Antequera,F. and Bird,A. (1993) Number of CpG islands and genes in human and mouse. *Proc. Natl. Acad. Sci. USA* **90**, 11995-11999.

Anziano,P.Q., Perlman,P.S., Lang,B.F., Wolf,K. (1983) The mitochondrial genome of the fission yeast *Schizosaccharomyces pombe*. I. Isolation and physical mapping of mitochondrial DNA. *Curr Genet.*, **7**, 273-284.

Barnitz,J.T., Cramer,J.H., Rownd,R.H., Cooley,L. and Soll,D. (1982) Arrangement of the ribosomal RNA genes in *Schizosaccharomyces pombe*. *FEBS Letts.*, **143**, 129-132.

Bartlett, R.C. (1991) PhD thesis, Oxford University.

Basi,G., Schmid,E. and Maundrell,K. (1993) TATA box mutations in the *Schizosaccharomyces pombe nmt1* promoter affect transcription efficiency but not the transcription start point or thiamine repressibility. *Gene*, **123**, 131-136.

Beach,D. (1983) Cell type switching by DNA transposition in fission yeast. *Nature*, **305**, 682-688.

Beach,D.H. and Klar,A.J.S. (1984) Rearrangements of the transposable mating-type cassettes of fission yeast. *EMBO J.*, **3**, 603-610.

Beach,D., Rodgers,L. and Gould,J. (1985) *RAN1*$^{+}$ controls the transition from mitotic division to meiosis in fission yeast. *Curr. Genet.*, **10**, 297-311.

Behe,M. and Felsenfeld,G. (1981) Effects of methylation on a synthetic polynucleotide: The B-Z transition in poly (dG-m^{5}dC). poly (dG-m^{5}dC). *Proc. Natl. Acad. Sci. USA* **78**, 1619-11623.

Bestor,T.H., Laudano,A., Mattaliano,R., Ingram,V. (1988) Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzyme is related to bacterial restriction methyltransferases. *J. Mol. Biol.*, **203**, 971-983.

Bestor,T.H. (1992) Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *EMBO J.*, **11**, 2611-2617.

Bird,A.P. (1978) Use of restriction enzymes to study eukaryotic DNA methylation: II. The symmetry of methylated sites supports semi-conservative copying of the methylation pattern. *J. Mol. Biol.*, **118**, 49-60.

Bird,A.P. and Southern,E.M. (1978) The methylation pattern in ribosomal DNA from *Xenopus inervis*. *J.Mol.Biol.*, **118**, 27-47.

Bird,A., Taggart,M. and Smith,B. (1979) Methylated and unmethylated DNA compartments in the sea urchin genome. *Cell*, **17**, 889-901.

Bird,A.P. (1980) DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res.*, **8**, 1499-1504.

Bird,A.P and Taggart,M.H. (1980) Variable patterns of total DNA and rDNA methylation in animals. *Nucleic Acids Res.*, **8**, 1485-1497.

Bird,A. (1986) CpG rich islands and the function of DNA methylation. *Nature*, **321**, 209-213.

Bird,A., Taggart,M., Nicholls,R. and Higgs,D. (1987) Non-methylated CpG rich islands at the human α-globin locus: implications for evolution of α-globin pseudogene. *EMBO J.*, **6**, 999-1004.

Boyes,J. and Bird,A. (1991) DNA methylation inhibits transcription indirectly via a methyl CpG binding protein. *EMBO J.*, **11**, 327-333

Broach,J.R., Strathern,J.N. and Hicks,J.B. (1979) Transformation in yeast: Development of a hybrid cloning vector and isolation of the *CAN1* gene. *Gene*, **8**, 121-131.

Brown,T.C. and Jircny,J. (1987) A specific mismatch repair event protects mammalian cells from loss of 5-methylcytosine. *Cell*, **50**, 945-950.

Buschhausen,G., Wittig,B., Graessman,M. and Graessman,H. (1987) Chromatin structure is required to block transcription of the methylated herpes simplex virus thymidine kinase gene. *Proc. Natl. Acad. Sci. USA* **84**, 1177-1181.

Cambereri,E.B., Jensen,B.C., Schabtach,E. and Selker,E.U. (1989) Repeat-induced G-C to A-T mutations in *Neurospora*. *Science*, **244**, 1571-1575.

Caserta,M., Zacharis,W., Nwankwo,D., Wilson,G.G. and Wells,R.D. (1987) Cloning, sequencing, *in vivo* promoter mapping and expression in *Escherichia coli* of the gene for the *HhaI* methyltransferase. *J. Biol. Chem.*, **262**, 4770-4777.

Cedar,H. (1988) DNA methylation and gene activity. *Cell*, **53**, 3-4.

Chen,L., MacMillan,A.M., Chang,W., Ezaz-Nikpay,K., Lane,W.S. and Verdine,G.L. (1991) Direct identification of the active-site nucleophile in a DNA (cytosine-5)-methyltransferase. *Biochemistry*, **30**, 11018-11025.

Chen,L., MacMillan,A.M. and Verdine,G.L. (1993) Mutational separation of DNA binding from catalysis in a DNA cytosine methyltransferase. *J. Am. Chem. Soc.*, **115**, 5318-5319.

Cheng,X., Kumar,S., Posfai,J., Pflugrath,J.W. and Roberts,R.J. (1993) Crystal structure of the HhaI DNA methyltransferase complexed with S-Adenosyl-L-methionine. *Cell*, **74**, 299-307.

Christman,J.K., Schneiderman,N. and Acs,G. (1985) Formation of highly stable complexes between 5-azacytidine substituted DNA and specific non-histone nuclear proteins. *J. Biol. Chem.*, **260**, 4059-4067.

Church,G. and Gilbert,W. (1984) Genomic sequencing. *Proc. Natl. Acad. Sci.* **81**, 1991-1995.

Collins,M. and Myers,R.M. (1987) Alterations in DNA helix stability due to base modifications can be evaluated using denaturing gradient gel electrophoresis. *J. Mol. Biol.*, **198**, 737-744.

Cooper,D., Taggart,M. and Bird,A. (1983) Unmethylated domains in vertebrate DNA. *Nucleic Acids Res.*, **11**, 647-658.

Coulondre,C., Miller,J., Farabaugh,P. and Gilbert,W. (1978) Molecular basis of base substitution hotspots in *Escherichia coli*. *Nature*, **274**, 775-780.

Crain,P.F. and McCloskey,J.A. (1983) Analysis of modified bases in DNA by stable isotope dilution gas chromatography-mass spectrometry: 5-methylcytosine. *Anal. Biochem.*, **132**, 124-131.

Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P. (1994) Purification of CpG islands using a methylated DNA binding column. *Nature Genetics*, **6**, 236-244.

Decaris,B., Lefort,C., Francou,F. and Rizet,G. (1979) Unstable ascospore colour mutants of *Ascobolus*. *Mol. Gen. Genet.*, **170**, 191-202.

Dennis,E. and Brettell,R. (1990) DNA methylation of maize transposable elements is correlated with activity. *Phil. Trans. R. Soc. Lond.*, **326**, 217-229.

Dila,D., Sutherland,E., Moran,L., Slatko,B. and Raleigh,E.A. (1990) Genetic and sequence organization of the McrBC locus of *Escherichia coli* K-12. *J. of Bacteriol.*, **172**, 4888-4900.

Eden,S. and Cedar,H. (1994) Role of DNA methylation in the regulation of transcription. *Curr. opin. genet. devel.*, **4**, 255-259.

Efstratiadis,A. (1994) Parental imprinting of autosomal mammalian genes. *Curr. opin. genet. devel.*, **4**, 265-280.

Egel,R. (1977) Frequency of mating type switching in homothallic fission yeast. *Nature*, **266**, 172-174.

Egel,R. (1984) Two tightly linked silent cassettes in the mating type of *Schizosaccharomyces pombe*. *Curr. Genet.* **8**, 199-203.

Egel,R., Beach,D.H. and Klar,A.J.S. (1984) Genes required for initiation and resolution steps of mating-type switching in fission yeast. *Proc. Natl. Acad. Sci. USA* **81**, 3481-3485.

Ekwall,K., Nielson,O. and Ruusala,T. (1991) Repression of a mating type cassette in the fission yeast by four DNA elements. *Yeast*, **7**, 745-755.

Ekwall,K. and Ruusala,T. (1994) Mutations in *rik1*, *clr2*, *clr3* and *clr4* asymmetrically derepress the silent mating type loci in fission yeast. *Genetics*, **136**, 53-64.

Faugeron,G. Rhounim,L. and Rossignol,J-L. (1990) How does the cell count the number of ectopic copies of a gene in the premeiotic inactivation process acting in *Ascobulus immersus*? *Genetics*, **124**, 585-591.

Feinberg,A.P. and Vogelstein,B. (1983) A technique for radiolabelling restriction endonuclease fragments to high specific activity. *Anal. Biochem.*, **1**, 6-13.

Finnegan,E.J. and Dennis,E.S. (1993) Isolation and identification by sequence homology of a putative cytosine methyltransferase from *Arabidopsis thaliana*. *Nucleic Acids Res.*, **21**, 2383-2388.

Foss,H.M., Roberts,C.J., Claeys,K.M. and Selker,E.U. (1993) Abnormal chromosome behaviour in *Neurospora* mutants defective in DNA methylation. *Science*, **262**, 1737-1741.

Freedman,T. and Pukkila,P.J. (1993) *De novo* methylation of repeated sequences in *Coprinus cinereus*. *Genetics*, **135**, 357-366.

Friedman,S. and Ansari,N. (1992) Binding of the *EcoRII* methyltransferase to 5-fluorocytosine-containing DNA: isolation of a bound protein. *Nucleic Acids Res.*, **20**, 3241-3248.

Fritzsche,E., Hayatsu,H., Igloi,G.L., Iida,S. and Kossel,H. (1987) The use of permanganate as a sequencing reagent for identification of 5-methylcytosine residues in DNA. *Nucleic Acids Res.*, **15**, 5517-5528.

Frommer,M., McDonald,L.E., Millar,D.S., Collis,C.M., Watt,F., Grigg,G.W., Molloy,P.L. and Paul,C.L. (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc. Natl. Acad. Sci. USA* **89**, 1827-1831.

Gadriner-garden,M. and Frommer,M. (1987) CpG islands in vertebrate genomes. *J. Mol. Biol.*, **196**, 261-282.

Gautsch,J.W. and Wilson,M.C. (1983) Delayed *de novo* methylation in tetracarcinoma suggests additional tissue specific mechanisms for controlling gene expression. *Nature*, **301**, 32-37.

Golberg,D.A. (1980) Isolation and partial characterization of the *Drosophila* alcohol dehydrogenase gene. *Proc. Natl. Acad. Sci. USA* **77**, 5794-5802.

Goyon,C. and Faugeron,G. (1989) Targeted transformation of *Ascobulus immersus* and *de novo* methylation of the resulting duplicated DNA sequences. *Mol. Cell. Biol.*, **9**, 2818-2827.

Goyon,C., Nogueira,T.I.V. and Faugeron,G. (1994) Perpetuation of cytosine methylation in *Ascobulus immersus* implies a novel type of maintenance methylase. *J. Mol. Biol.*, **240**, 42-51.

Grigg,G. and Clark,S. (1994) Sequencing 5-methylcytosine residues in genomic DNA. *Bioessays*, **16**, 431-436.

Grimm,C., Kohli,J., Murray,J. and Maundrell,K. (1988) Genetic engineering of *Schizosaccharomyces pombe*: A system for gene disruption and replacement using the *ura4* gene as a selectable marker. *Mol. Gen. Genet.*, **215**, 81-86.

Gruenbaum,Y., Naveh-Many, T., Cedar,H. and Razin,A. (1981a) Sequence specificity of methylation in higher plant DNA. *Nature*, **292**, 860-862.

Gruenbaum,Y., Stein,R., Cedar,H. and Razin,A. (1981b) Methylation of CpG sequences in eukaryotic DNA. *FEBS Letts.*, **124**, 67-71.

Gutz,H., Heslot,H., Leupold,U. and Loprieno,N. (1974) *Schizosaccharomyces pombe*. In: *Handbook of genetics*, Volume 1 (R.C.King Ed). Plenum Publishing Corp., New York.

Gutz,H. and Schmidt,H. (1985) Switching genes in *Schizosaccharomyces pombe*. *Curr. Genet.*, **9**, 325-331.

Haber,J.E., Thorburn,P.C. and Rogers,D. (1984) Meiotic and mitotic behaviour of dicentric chromosomes in *Saccharomyces cerevisiae*. *Genetics*, **106**, 185-205.

Hanck,T., Schmidt,S. and Fritz,H-J. (1993) Sequence-specific and mechanism-based cross-linking of Dcm DNA cytosine-C^5 methyltransferase and *E. coli* K-12 to synthetic oligonucleotides containing 5-fluoro-2'deoxycytidine. *Nucleic Acids Res.*, **21**, 303-309.

Harlow,E. and Lane,D. (1988) Antibodies-A laboratory manual. Cold Spring Harbor publications.

Henikoff,S. (1984) Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. *Gene*, **28**, 351-359.

Hieter,P., Mann,C., Snyder,M. and Davis,R.W. (1985) Mitotic stability of yeast chromosomes: A colony color assay that measures nondisjunction and chromosome loss. *Cell*, **40**, 381-392.

Hindley,J., Phear,G., Stein,M. and Beach,D. (1987) *suc1*$^+$ encodes a predicted 13 kilodalton protein that is essential for cell viability and is directly involved in the division cycle of *Schizosaccharomyces pombe*. *Mol. Cell. Biol.*, **7**, 504-511.

Hoheisel,J.D., Maier,E., Mott,R., McCarthy,L., Grigoriev,A.V., Schalkwyk,L.C., Nizetic,D., Francis,F. and Lehrach,H. (1993) High resolution cosmid and P1 maps spanning the 14Mb genome of the fission yeast *S. pombe*. *Cell*, **73**, 109-120.

Humbelin,M., Suri,B., Rao,D.N., Hornby,D.P., Eberle,H. (1988) Type III DNA restriction and modification systems *EcoPI* and *EcoP15*: Nucleotide sequence of the *EcoP1* operon, the *EcoP15* modification gene and some *EcoPI* modification mutants. *J. Mol. Biol.*, **200**, 23-29.

Humphrey,T., Sadhale,P., Platt. and Proudfoot,N. (1991) Homologous mRNA 3' end formation in fission and budding yeast. *EMBO.J.*, **10**, 3503-3511.

Iguchi-Ariga,S.M.M. and Schaffner,W. (1989) CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. *Genes and Dev.*, **3**, 612-619.

Iino,Y. and Yamamoto,M. (1985) Mutants of *Schizosaccharomyces pombe* which sporulate in the haploid state. *Mol. Gen. Genet.*, **198**, 416-421.

Ingrosso,D., Fowler, Fowler,A.V., Bleibaum,J. and Clarke,S. (1989) Sequence of the D-aspartyl L-isoaspartyl protein methyltransferase from human erythrocytes. Common sequence motifs for protein, DNA, RNA and small molecule S-adenosylmethionine-dependent methyltransferases. *J. Biol. Chem.*, **264**, 20130-20139.

Jaenisch,R. (1982) Retroviruses and embryogenesis: *de novo* methylation activity involved in gene expression. *CSH Symp. Quant. Biol.*, **47**, 611.

Jaworski,A., Zacharias,W., Hsieh,W-T., Blaho,J.A., Larson,J.E. and Wells,R.D. (1988) *In vivo* existence of left-handed DNA. *Gene*, **74**, 215-220.

Jones,P. (1984) Gene activation by 5-azacytidine. In DNA methylation: Biochemistry and biological significance. Edited by Razin,A. Cedar,H. and Riggs,A.D. New York: Springer Verlag. 165-187.

Jones,M., Wagner,R. and Radman,M. (1987) Mismatch repair of deaminated 5-methylcytosine. *J. Mol. Biol.* **194**, 155-159.

Jupe,E.R., Magill,J.M. and Magill,C.W. (1986) Stage-specific DNA methylation in a fungal pathogen. *J. Bacteriol.*, **165**, 420-423.

Kafri,T., Ariel,M., Brandeis,M., Shemer,R., Urven,L., McCarrey,J., Cedar,H. and Razin,A. (1992) Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev.*, **6**, 705-714.

Kaufer,N.F., Simanis,V. and Nurse,P. (1985) Fission yeast *Schizosaccharomyces pombe* correctly excises a mammalian RNA transcript intervening sequence, *Nature*, **318**, 78-80.

Kelley,D., Pollok,B.A., Atchison,M.L. and Perry,R.P. (1988) The coupling between enhancer activity and hypomethylation of κ-immunoglobulin genes is developmentally regulated. *Mol. Cell. Biol.*, **8**, 930-937.

Keshet,I., Lieman-Hurwitz,J. and Cedar,H. (1986) DNA methylation affects the formation of active chromatin. *Cell*, **44**, 535-543.

Klar,A. (1987) Differential parental DNA strands confer developmental asymmetry on daughter cells in fission yeast. *Nature*, **326**, 466-470.

Klimasauskas,S., Timinskas,A., Menkevicius,S., Butkiene,D., Butkus,V. and Janulaktis,A.A. (1989) Sequence motifs characteristic of DNA [cytosine-N4] methyltransferases: similarity to adenine and cytosine-C5 DNA methyltransferases. *Nucleic Acids Res.*, **17**, 9823-9832.

Klimasauskas,S., Nelson,J.L. and Roberts,R.J. (1991) The sequence specificity domain of cytosine-C5 methylases. *Nucleic Acids Res.*, **19**, 6183-6190.

Klimasauskas,S., Kumar,S., Roberts,R.J. and Cheng,X. (1994) *HhaI* methyltransferase flips its target base out of the DNA helix. *Cell*, **76**, 357-369.

Kohli,J., Hottiger,H., Munz,P., Strauss,A. and Thuriaux,P. (1977) Genetic mapping in *Schizosaccharomyces pombe* by mitotic and meiotic analysis, and induced haploidization. *Genetics*, **87**, 423-471.

Kolodziej,P. and Young,R.A. (1991) Epitope tagging and protein surveillance. *Methods in Enzymol.*, **194**, 508-519.
Kovesdi,I., Reichel,R. and Nevins,J.R. (1987) Role of adenovirus E2 promoter binding factor in E1A-mediated co-ordinate gene control. *Proc. Natl. Acad. Sci.*, **84**, 2180.

Kozak,M. (1983) Comparison of initiation of protein synthesis in prokaryotes, eukaryotes and organelles. *Microbial reviews*, **47**, 1-45.

Kruger,D., Barcak,G.J., Reuter,M. and Smith,H.O. (1988) *EcoR*II can be activated to cleave refractory DNA recognition sites. *Nucleic Acids Res.*, **16**, 3997-4008.

Kuo,K.C., McCure,R.A. and Gehrke,C.W. (1980) Quantitative reversed-phase high performance liquid chromatographic determination of major and modified deoxyribonucleosides in DNA. *Nucleic Acids Res.*, **8**, 4763-4776.

Laemmli,U.K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T_4. *Nature*, **227**, 680.

Leupold,U. (1950) Die vererbung von homoyhallie und heterothallie bie *Schizosaccharomyces pombe. C.R.Trav.Lab.Carlesberg,Ser. Physiol.* **24**, 381-480.

Leupold,U. (1970) *Methods Cell Physiol.*, **4**, 169-177. Genetical methods for *Schizosaccharomyces pombe.*

Levin,H.L., Weaver,D.C. and Boeke,J.D. (1990) Two related families of retrotransposons from *Schizosaccharomyces pombe. Mol. Cell. Biol.*, **10**, 6791-6798.

Lewis,J.D., Meehan,R.R., Henzel,W.J., Mauer-Fogy,I., Jeppesen,P., Klein,F. and Bird,A. (1992) Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell*, **69**, 905-914.

Lewis,J. and Bird,A. (1991) DNA methylation and chromatin structure. *FEBS Lett.*, **285**, 155-159.

Li,R. and Murray,A.W. (1991) Feedback control of mitosis in budding yeast. *Cell*, **66**, 519-531.

Li,E., Bestor,T.H. and Jaenisch,R. (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, **69**, 915-926.

Li,R., Havel,C., Watson,J.A. and Murray,A.W. (1993) The mitotic feedback control gene *MAD2* encodes the alpha subunit of a prenyltransferase. *Nature*, **366**, 82.

Li,R., Havel,C., Watson,J.A. and Murray,A.W. (1994) *Nature*, 371, 438.

Lin,P.M., Lee,C.H. and Roberts,R.J. (1989) Cloning and characterization of the genes encodong the *MspI* restriction modification system. *Nucleic Acids Res.*, **17**, 3001-3011.

Lipman,D.J. and Pearson,W.R. (1985) Rapid and sensitive protein similarity searches. *Science*, **227**, 1435-1441.

Little,P.F.R. and Cross,S.H. (1985) A cosmid vector that facilitates restriction enzyme mapping. *Proc. Natl. Acad. Sci. USA* **82**, 3159-3163.

Lorentz,A., Heim,L. and Schmidt,H. (1992) The switching gene *swi6* affects recombination and gene expression in the mating-type region of *Schizosaccharomyces pombe*. *Mol. Gen. Genet.*, **233**, 436-442.

Lorentz,A., Osterman,K., Fleck,O. and Schmidt,H. (1994) Switching gene, *swi6*, involved in repression of silent mating-type loci in fission yeast encodes a homologue of chromatin-associated proteins from *Drosophila* and mammals. *Gene*, **142**, 139-143.

Macleod,D., Charlton,J., Mullins,J. and Bird,A.P. (1994) Sp1 sites in the mouse *aprt* gene promoter are required to prevent methylation of the CpG island. *Genes and Development*, **8**, 2282-2292.

Maden,B.E.H. (1988) Locations of methyl groups in 28S rRNA of *Xenopus laevis*. *J. Mol. Biol.*, **201**, 289-314.

Maier, E., Hoheisel,J.D., McCarthy, L., Mott,R., Grigorev,A.V., Monaco,A.P., Larin,Z. and Lehrach,H. (1992) Complete coverage of the *Schizosaccharomyces pombe* genome in yeast artificial chromosomes. *Nature Genetics*, **1**, 273-277.

Maundrell,K. (1990) *nmt1* of fission yeast. *J. of Biol. Chem.*, **265**, 10857-10864.

Maundrell,K. (1993) Thiamine repressible expression vectors pREP and pRIP for fission yeast. *Gene*, **123**, 127-130.

Maxam,A.M. and Gilbert,W. (1980) Sequencing end labelled DNA with base specific chemical cleavages. *Meth. Enzymol.*, **65**, 499-560.

McLeod,M. and Beach,D. (1986) Homology between the *ran1*$^+$ gene of fission yeast and protein kinases. *EMBO J.*, **5**, 3665-3671.

McLeod,M., Stein,M. and Beach,D. (1987) The product of the *mei3*$^+$ gene expressed under control of the mating-type locus, induces meiosis and sporulation in fission yeast. *EMBO J.*, **6**, 729-736.

McLeod,M. and Beach,D. (1988) A specific inhibitor of the *ran1*$^+$ protein kinase regulates entry into meiosis in *Schizosaccharomyces pombe*. *Nature*, **332**, 509-514.

Meehan,R.R., Lewis,J.D., McKay,S., Kleiner,E.L. and Bird,A.P. (1989) Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. *Cell*, **58**, 499-507.

Meehan,R.R., Lewis,J.D. and Bird,A.P. (1992) Characterization of MeCP2, a vertebrate DNA binding protein that binds methylated DNA. *Nucleic Acids Res.*, **20**, 5085-5092.

Messer,W. and Noyer-Weidner,M. (1988) Timing and targeting: The biological functions of Dam methylation in *E. coli*. *Cell*, **54**, 735-737.

Meyer,P., Niedenhof,I. and ten Lohuis,M. (1994) Evidence for cytosine methylation of non-symmetrical sequences in transgenic *Petunia hybrida*. *EMBO J.*, 13, 2084-2088.

Mi,S. and Roberts,R.J. (1992) How M.*Msp*I and M.*Hpa*II decide which base to methylate. *Nucleic Acids Res.*, **20**, 4811-4816.

Mi,S. and Roberts,R.J. (1993) The DNA binding affinity of the HhaI methylase is increased by a single amino acid substitution in the catalytic centre. *Nucleic Acids Res.*, **21**, 2459-2464.

Mitchison,J.M. (1970) Physiological and cytological methods for *Schizosaccharomyces pombe*. *Methods Cell Physiol.*, **4**, 131-165.

Modrich,P. (1991) Mechanisms and biological effects of mismatch repair. *Annu. Rev. Genet.*, **25**, 229-253.

Monk, M., (1987) Genomic imprinting-memories of mother and father. *Nature*, **328**, 203-204.

Monk,M., Boubelik,M. and Lehnert,S. (1987) Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development*, **99**, 371-382.

Moreno,S., Klar,A. and Nurse.P. (1991) Molecular genetic analysis of the fission yeast *Schizosaccharomyces pombe*. *Methods Enzymol.*, **194**, 795-823.

Mueller,P.R. and Wold,B. (1989) *In vivo* footprinting at a developmentally regulated enhancer. *Science*, **246**, 780-786.

Murakami,S., Matsumoto,T., Niwa,O. and Yanagida,M. (1991) Structure of the fission yeast centromere cen3: Direct analysis of the reiterated region. *Chromosoma*, **101**, 214-221.

Nicolas,A., Hamza,H., Mekki-Berrada,A., Kalogeropoulos,A. and Rossignol,J.-L. (1987) Premeiotic and meiotic instability generates numerous b2 mutation derivatives in *Ascobulus*. *Genetics*, **116**, 33-43.

Nimmo,E.R., Cranston,G. and Allshire,R.C. (1994) Telomere-associated chromosome breakage in fission yeast results in variegated expression of adjacent genes. *EMBO J.*, **13**, 3801-3811.

Niwa,O., Matsumoto,T. and Yanagida,M. (1986) Construction of a mini-chromosome by deletion of its mitotic and meiotic behaviour in fission yeast. *Mol. Gen. Genet.*, **203**, 397-405.

Noller,H.F. (1984) Structure of ribosomal RNA. *Annu. Rev. Biochem.*, **53**, 119-162.

Nurse,P. (1975) Genetic control of cell size at cell division in yeast. *Nature*, **256**, 547-551.

Nurse,P. (1985) Mutants of the fission yeast *Schizosaccharomyces pombe* which alter the shift between cell proliferation and sporulation. *Mol. Gen. Genet.*, **198**, 497-502.

Osterman,D.G., DePillis,G.D., Wu,J.C., Matsuda,A. and Santi,D.V. (1988) 5-fluorocytosine in DNA is a mechanism based inhibitor of *HhaI* methylase. *Biochemistry*, **27**, 5204-5210.

Pein,C.D., Reuter,M., Meisel,A., Cech,D. and Kruger,D.H. (1991) Activation of restriction endonuclease *Eco*RII does not depend on the cleavage of stimulator DNA. *Nucleic Acids Res.*, 19, 51389-5142.

Perkins,D.D., Metzenberg,R.L., Raju,N.B., Selker,E.U. and Barry,G. (1986) Reversal of a *Neurospora* translocation by crossing over involving displaced rDNA and methylation of the rDNA segments that result from recombination. *Genetics*, **114**, 791-

Pfeifer,G.P., Steigerwald,S.D., Mueller,P.R., Wold,B. and Riggs,A.D. (1989) Genomic sequencing and methylation analysis by ligation mediated PCR. *Science*, 246, 810-813.

Posfai,J., Bhagwat,A.S., Posfai,G. and Roberts,R.J. (1989) Predictive motifs derived from cytosine methyltransferases. *Nucleic acids Res.*, 17, 2451-2435.

Prentice, H.L. (1991) High efficiency transformation of *Schizosaccharomyces pombe*. *Nucleic acids Res.*, 20, 621.

Proffitt,J.H., Davie,J.R., Swinton,D. and Hattman,S. (1984) 5-methylcytosine is not detectable in *Saccharomyces cerevisiae* DNA. *Mol. Cell. Biol.*, 4, 985-988.

Pruitt,R.E. and Meyerowitz,E.M. (1986) Characterization of the genome of *Arabidopsis thaliana*. *J. Mol. Biol.*, 187, 169.

Razin,A. and Sedat,J. (1977) *Anal. Biochem.*, 77, 370-377.

Razin,A. and Cedar,H. (1991) DNA methylation and gene expression. *Microbial Rev.*, 55, 451-458.

Reddy,M.V., Gupta,R.C. and Randerath,K. (1981) ^{32}-P base analysis of DNA. *Anal. Biochem.*, 117, 271-279.

Renbaum,P., Abrahamove,D., Fainsod,A., Wilson,G.G., Rottem,S. and Razin,A. (1990) Cloning, characterization and expression in *Escherichia coli* of the gene coding for the CpG DNA methylase from *Spiroplasma sp.* strain MQ1 (M.SssI). *Nucleic Acids Res.*, 18, 1145-1152.

Riggs,A.D. and Pfeifer,G.P. (1992) X-chromosome inactivation and cell memory. *Trends Genet.*, 8, 169-174.

Rothnie,H.M., McCurrack,K.J., Glover,L.A. and Hardman,N. (1991) Retrotransposon-like nature of TP1 elements-implications for the organization of highly repetitive, hypermethylated DNA in the genome of *Physarum polycephalum*. *Nucleic Acids Res.*, 19, 279-286.

Rothstein,R.J. (1983) One step gene disruption in yeast. In: *Methods in Enzymology* (R.Wu, L.Grossman and K.Moldave eds), Academic Press, 101, 202-211.

Russell,P.J., Rodland,K.D., Cutler,J.E., Rachlin,E.M. and McCloskey,J.A. (1985) In Molecular Genetics of filamentous fungi, edited by W. Timberlake. Alan R. Liss. New York. pp 321-332.

Russell,P.J., Welsch,J.A., Rachlin,E.M. and McCloskey,J.A. (1987a) Different levels of DNA methylation in yeast and mycelial forms of *Candida albicans*. *J. Bacteriol.*, 169, 4393-4395.

Russell,P.J., Rodland,K.D., Rachlin,E.M. and McCloskey,J.A. (1987b) Differential DNA methylation during the vegetative life cycle of *Neurospora crassa*. *J. Bacteriol.*, 169, 2902-2905.

Russell,P. (1989) Gene cloning and expression in fission yeast. In Nasim,A., P. Young and B. Johnson, Academic Press, USA. pp244-271.

Saluz,H.P. and Jost,J.P. (1986) Optimized genomic sequencing as a tool for the study of cytosine methylation in the regulatory region of the chicken vitellogenin II gene. *Gene*, **42**, 151-157.

Sambrook,J., Fritsch,E.F. and Maniatis, T. (1989) Molecular cloning: A laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.

Santi,D.V. and Danenberg,P.V. (1984) In folates and pterins, volume I, R.L. Blakely and S.J. Benkovic, eds. (New York: Wiley-Interscience). pp343-396.

Santi,D.V., Norment,A. and Garrett,C.E. (1984) Covalent bond formation between a DNA-cytosine methyltransferase and DNA containing 5-azacytidine. *Proc. Natl. Acad. Sci USA* **81**, 6993-6997.

Santi,D.V. and Hardy,L.W. (1987) Catalytic mechanism and inhibition of tRNA (uracil-5-) methyltransferase: Evidence for covalent catalysis. *Biochemistry*, **26**, 8599-8606.

Sasaki,H., Jones,P.A., Chaillet,J.R., Fergusson-Smith,A.C., Barton,S.C., Reik,W and Surani,M.A. (1992) Parental imprinting: Potentially active chromatin of the repressed maternal allele of the mouse insulin-like growth factor II Igf2 gene. *Genes and deveolpment*, **6**, 1843-1856.

Scheidt,G., Weber,H., Graessmann,M. and Graessmann,A. (1994) Are there two methyltransferase gene families in plant cells? A new potential methyltransferase gene isolated from an *Arabidopsis thaliana* genomic library. *Nucleic Acids Res.*, **22**, 953-958.

Schein,C.H. (1989) Production of soluble recombinant proteins in bacteria. *Biotechnology*, **7**, 1141-1149.

Schmidt,H., Kapitaza,P. and Gutz,H. (1987) Switching genes in *Schizosaccharomyces pombe*: their influence on cell viability and recombination. *Curr. Gent.*, **11**, 303-308.

Selker,E.U., Jensen,B.C. and Richardson,G.A. (1987) A portable signal causing faithful DNA methylation *de novo* in *Neurospora crassa*. *Science*, **238**, 48-53.

Selker,E.U. and Garrett,P.W. (1988) DNA sequence duplications trigger gene inactivation in *Neurospora*. *Proc. Natl. Acad. Sci. USA*. **85**, 6870-6874.

Selker,E.U., Cambareri,E.B., Jensen,B.C. and Haack,K.R. (1987) Rearrangement of duplicated DNA in specialized cells of *Neurospora*. *Cell*, **51**, 741-752.

Selker,E.U. (1990) DNA methylation and chromatin structure: a view from below. *Trends Biochem. Sci.*, **15**, 103-107.

Selker,E.U., Fritz,D.Y. and Singer,M.J. (1993) Dense nonsymmetrical DNA methylation resulting from repeat-induced point mutations in *Neurospora*. *Science*, **262**, 1724-1728.

Shapiro,H.S. (1975) Content of 6-methylaminopurine and 5-methylcytosine in DNA. CRC Press. Boca Raton, FL.

Simpson,V.J., Johnson,T.E. and Hammen,R.F. (1986) *Caenorhabditis elegans* DNA does not contain 5-methylcytosine at any time during development or aging. *Nucleic Acids Res.*, **14**, 6711-6719.

Singer,J., Stellwagen,R.H., Roberts-Ems,J.. and Riggs,A.D. (1977) 5-methylcytosine content of rat hepatoma DNA substituted with bromodeoxyuridine. *J. Biol. Chem.*, **252**, 5509-5513.

Singer-Sam,J., Yang,T.P., Mori,N., Tanguay,A.L., LeBon,J.M., Flores,J.C. and Riggs,A.D. (1990) DNA methylation in the 5' region of the mouse PGK-1 gene and a quantitative PCR assay for methylation. In: UCLA New Series Symposia on Molecular and Cellular Biology: Nucleic Acid Methylation (ed. G. Clawson, D. Willis, A. Weissback and P. Jones). pp285-298. Wiley-Liss: New York.

Smith,C.L., Matsumoto,T., Niwas,O., Kleo,S., Binfan,J., Yanagida,M. and Cantor,C.R. (1987) An electrophoretic karyoptype for *S. pombe* by pulse field gel electrophoresis. *Nucleic Acids Res.*, **15**, 4481-4489.

Smith,S.S., Kaplan,B.E., Sowers,L.C. and Newman,E.M. (1992) Mechanisms of human methyl-directed DNA methyltransferase and the fidelity of cytosine methylation. *Proc. Natl. Acad. Sci. USA.* **89**, 4744-4748.

Som,S., Bhagwat,A.S. and Friedman,S. (1987) Nucleotide-sequence and expression of the gene encoding the *EcoRII* modification enzyme. *Nucleic Acids Res.*, **15**, 313-332.

Southern,E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.*, **98**, 503-517.

Specht,C.A., Novotny,C.P. and Ullrich,R.C. (1984) Strain specific differences in ribosomal DNA from the fungus *Schizophyllum commune*. *Curr. Genet.*, **8**, 219-222.

Stewart,C.L., Stuhlmann,H., Jahner,D. and Jaenisch,R. (1982) *De novo* methylation, expression and infectivity of retroviral genomes introduced into embryonal carcinoma cells. *Proc. Natl. Acad. Sci. USA.* **79**, 4098-4102.

Studier,F.W., Rosenberg,A.H., Dunn,J.J. and Dubendorff,J.W. (1990) Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.*, **185**, 60-89.

Sugawara,N. and Szostak,J.W. (1986) *Yeast*, **2**, (Suppl) 373.

Sugisaki,H., Yamamoto,K. and Takanami,M. (1991) The *HgaI* restriction-modification system contains two cytosine methylase genes responsible for modification of different DNA strands. *J. Biol. Chem.*, **266**, 13952-13957.

Sutherland,E., Coe,L. and Raleigh,E.A. (1992) McrBC: a multisubunit GTP-dependent restriction endonuclease. *J. Mol. Biol.*, **225**, 327-348.

Sutter,D. and Doerfler,W. (1980) Methylation of integrated adenovirus type 12 DNA sequences in transformed cells is inversely correlated with viral gene expression. *Proc. Natl. Acad. Sci. USA* **77**, 253-256.

Sved,J. and Bird,A. (1990) The expected equilibrium of the CpG dinucleotide in vertebrate genomes under a mutation model. *Proc. Natl. Acad. Sci. USA* **87**, 4692-4696.

Tasheva,E.S. and Roufa,D.J. (1994) Densely methylated DNA islands in mammalian chromosomal replication origins. *Mol. Cell. Biol.*, **14**, 5636-5644.

Taylor,C., Ford,K., Connolly,B.A. and Hornby,D.P. (1993) Determination of the order of substrate addition to *MspI* DNA methyltransferase using a novel mechanism based inhibitor. *Biochem. J.*, **291**, 493-504.

Thon,G and Klar,A.J.S. (1992) The *clr1* locus regulates the expression of the cryptic mating-type loci of fission yeast. *Genetics*, **131**, 287-296.

Thon,G., Cohen,A. and Klar,A.J. (1994) Three additional linkage groups that repress transcription and meiotic recombination in the mating-type region of *Schizosaccharomyces pombe*. *Genetics*, **138**, 29-38.

Tommasino,M. and Maundrell,K. (1991) Uptake of thiamine by *Schizosaccharomyces pombe* and its effect on a transcriptional regulator of thiamine-sensitive genes. *Curr. Genet.*, 20, 63-66.

Urieli-Shoval,S., Gruenbaum,Y., Sedat,J. and Razin,A. (1982) The absence of detectable methylated bases in *Drosophila melanogaster* DNA. *Febs Letts.*, **146**, 148-152.

Van Etten,J.L., Xia,Y. and Burbank,D.E. (1989) Viruses infecting a eukaryotic green alga are a new source of DNA methyltransferases and DNA site specific endonuclease. *J. Cell Biochem. suppl.*, 199.

Vardimon,L., Kressman,A., Cedar,H., Maechler,M. and Doerfler,W. (1982) Expression of a cloned adenovirus gene is inhibited by *in vitro* methylation. *Proc. Natl. Acad. Sci. USA* **79**, 1073.

Viera,K. and Messing,J. (1982) The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene*, **19**, 259-268.

Vongs,A., Kakutani,J., Martienssen,R.A. and Richards,E.J. (1993) *Arabidopsis thaliana* DNA methyltransferase mutants. *Science*, **260**, 1926-1928.

Waalwijk,C. and Flavell,R.A. (1978) *MspI*, an isoschizomer of *HpaII* which cleaves both unmethylated and methylated *HpaII* sites. *Nucleic Acid Res.*, **5**, 3231-3236.

Watt,F. and Molloy,P.L. (1988) Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression and the adenovirus major late promoter. *Genes and development*, **2**, 1136-1143.

Wilke,K., Rauhut,E., Noyer-Weidener,M., Lauster,R., Pawlek,B., Behrens,B. and Trautner,T.A. (1988) Sequential order of target-recognizing domains in multispecific DNA-methyltransferases. *EMBO J.* **7**, 2601-2609.

Wilson,G.G. and Murray,N.E. (1991) Restriction and modification systems. *Annu. Rev. Genet.*, **25**, 585-627.

Wright,A., Maundrell,K., Heyer,W-D., Beach,D. and Nurse,P. (1986) Vectors for the construction of gene banks and integration of cloned genes in *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae*. *Plasmid*, **15**, 156-158.

Wu,J.C. and Santi,D.V. (1985) In Biochemistry and Biology of DNA methylation, G.I. Cantoni and A.Razin, eds (New York: Alan R. Liss Incorporated. pp119-129.

Wyszynski,M.W., Gabbara,S. and Bhagwat,A.S. (1992) Substitutions of a cysteine conserved among DNA cytosine methylases results in a variety of phenotypes. *Nucleic Acids Res.*, **20**, 319-326.

Yen,R-W,C., Vertino,P.M., Nelkin,B.D., Yu,J.J., El-Deiry,W., Cumaraswamy,A., Lennon,G.G.., Trask,B.J., Celano,P. and Baylin,S.B. (1992) Isolation and characterization of the cDNA encoding human DNA methyltransferase. *Nucleic Acids Res.*, **20**, 2287-2291.

Zaret,K.S. and Sherman,F. (1982) DNA sequence requirement for efficient transcription termination in yeast. *Cell*, **28**, 563-573.

Zhang,Y., Nelson,M. and Van Etten,J.L. (1992) A single amino acid change restores DNA cytosine methyltransferase activity in a cloned *chlorella* virus. *Nucleic Acids Res.*, **20**, 1637-1642.

Zolan,M.E. and Pukkila,P.J. (1985) DNA methylation in *Coprinus cinereus* in Molecular Genetics of Filamentous Fungi, edited by W. Timberlake. Alan R. Liss, New York. pp 333-344.

Zolan,M.E. and Pukkila,P. (1986) Inheritance of DNA methylation in *Coprinus cinereus*. *Mol. Cell. Biol.*, **6**, 195-200.

Appendix A-Oligonucleotides I

Oligonucleotides used for sequencing pBsup1-2 (see chapter 4).

A 5' AGGTTGAGGCCGTTGAGC 3'
B 5' CCGAAGCCTCGCACATAT 3'
C 5' CGGCTAGTCTTGCAACTG 3'
D 5' GGGTCCACCGCTTGTT 3'

Oligonucleotides used for the amplification of the *S. pombe* LTR

LTR-1 5' GTGCGAATTCGCGCTACGCAGTTTGGTATTTGA 3'
LTR-2 5' GTGCGAATTCGCGTCAGCAATACTACACTACGC 3'

PCR conditions (30 cycles): 1 minute 94ºC
1 minute 55ºC
1 minute 72ºC

Oligonucleotides containing the mating type locus (used in the *in vitro* m5C-MTase assays chapter 5).

mat1P-A 5' GTTTTTGTTTTTCATAAATTTTTTTTTTGTAATATAAATG 3'
mat1P-B 5' CATTTATATTACAAAAAAAAAATTTATGAAAAACAAAAAC 3'

mat1M-A 5' GTTTTTATTTATTTTCAATAATTTTTTTGTAATATAAATG 3'
mat1M-B 5' CATTTATATTACAAAAAAATTATTGAAAATAAATAAAAAC 3'

These oligonucleotides were mixed together in equal quantities, placed in boiling water for 2 minutes and then allowed to cool down at 37ºC to allow annealing to occur.

Appendix B-Oligonucleotides II

i Oligonucleotides for subcloning *cnd1* (*pmt1*) cDNA into pET6H

cnd15 5' CGG**CCATGG**GGATGCTTAGTACAAAAAGATTAC 3'
NcoI

cnd13 5' GCT**GGATCC**CCGCTTTTAGAAATTTAGAGGTTC 3'
BamHI

PCR conditions (30 cycles): 1 minute 94ºC
1 minute 55ºC
2 minutes 72ºC

ii Oligonucleotides for sequencing *cnd1* (*pmt1*) cDNA

cndA 5' CCGGCTGATATTGTGTGT 3'
cndB 5' AGTCCCAGTTGTCAACCT 3'
cndC 5' AGAGCAAGGCAGCTGAAG 3'
cndD 5' AATTTAGCGAGGTAGCGC 3'
cndE 5' GTTGCTTTACACGGGGAT 3'

861 (oligonucleotide complementary to the 3' end of the nmt1 promoter in pREP1)
5' GGAATCCTGGCATATCATCAATTG 3'

iii Oligonucleotides for subcloning *cnd1* (*pmt1*) cDNA into pREP1

Ndecnd5 5' GCGAC**CATATG**CTTAGTACAAAAAGATTACGGG 3'
NdeI

cnd13 as above

PCR conditions (30 cycles): 1 minute 94ºC
1 minute 55ºC
1 minute 72ºC

iv Oligonucleotides spanning the intron in *pmt1* for use in RT-PCR

RT5 5' TTACGGGTCCTGGAGCTATAT 3'
RT3 5' GACAACTGGGACTCATAGTCC 3'

PCR conditions (30 cycles); 1 minute 94ºC
1 minute 55ºC
2 minutes 72ºC

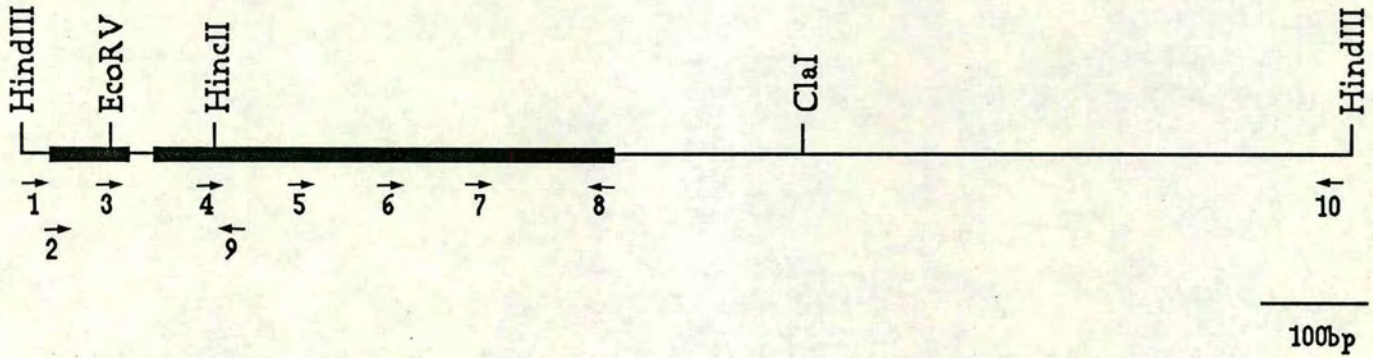
v Oligonucleotides to amplify the *pmt1::ura4* construct (see chapter 7)

5'kopmt 5' AGCTTAACAAATCCAAGCAA 3'
3'kopmt 5' GTAGCTAAAGCATTCAATAC 3'

PCR conditions: 1 minute 94ºC
1 minute 55ºC
1 minute 72ºC

This reaction was carried out using Thesit buffer (section 2.6).

The locations of all these oligonucleotides within the *pmt1*$^+$ genomic clone are shown in the diagram on the following page.



Appendix B- Restriction map of pmt1 genomic clone showing positions of oligonucleotides

pmt1 ORFs are shown as black boxes.

Oligonucleotides are as follows:

1 = 5'kopmt	6 = cndD
2 = Ndecnd5/ RT5/ cnd15	7 = cndE
3 = cndA	8 = cnd13
4 = cndB	9 = RT3
5 = cndC	10 = 3'kopmt